



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

– **TELECOM** ESCUELA
TÉCNICA **VLC** SUPERIOR
DE INGENIERÍA DE
TELECOMUNICACIÓN

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería de
Telecomunicación

Uso de IA en edición audiovisual y aplicación en redes
sociales

Trabajo Fin de Grado

Grado en Ingeniería de Tecnologías y Servicios de
Telecomunicación

AUTOR/A: Paineira Pérez, Jesse

Tutor/a: Sastre Martínez, Jorge

CURSO ACADÉMICO: 2023/2024



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

– **TELECOM** ESCUELA
TÉCNICA **VLC** SUPERIOR
DE INGENIERÍA DE
TELECOMUNICACIÓN

NO PONER PORTADA

Escuela Técnica Superior de Ingeniería de Telecomunicación
Universitat Politècnica de València
Edificio 4D. Camino de Vera, s/n, 46022 Valencia
Tel. +34 96 387 71 90, ext. 77190
www.etsit.upv.es

VLC/
CAMPUS
VALENCIA, INTERNATIONAL
CAMPUS OF EXCELLENCE



Resumen

En este trabajo se ha realizado una investigación sobre las diferentes aplicaciones para edición audiovisual basadas total o parcialmente en Inteligencia Artificial, analizando las fortalezas y debilidades de cada una de ellas y enfocando su uso hacia la generación de contenido en redes sociales de una manera rápida y eficiente. Entre las aplicaciones analizadas se encuentran herramientas para editar, mezclar y masterizar sonido, para optimizar y mejorar grabaciones de voz, para reducir ruido y sonidos no deseados, para resumir y editar vídeos obteniendo cortes con los mejores momentos y para optimizar dichos contenidos editados para su utilización en redes sociales dependiendo de la plataforma en la que se deseen publicar.

El objetivo es, una vez estudiadas las opciones, obtener un procedimiento estándar a seguir en la práctica cuando se tiene un contenido audiovisual en crudo (sin procesar) para editarlo y prepararlo de la manera más sencilla y rápida posible utilizando estas herramientas existentes.

Se plantea además una reflexión adicional sobre el uso y posibilidades de la Inteligencia Artificial en este sector en particular.

Resum

En este treball s'ha realitzat una investigació sobre les diferents aplicacions per a edició audiovisual basades totalment o parcialment en Intel·ligència Artificial, analitzant les fortaleres i debilitats de cadascuna d'elles i enfocant el seu ús cap a la generació de contingut en xarxes socials d'una manera ràpida i eficient. Entre les aplicacions analitzades es troben eines per a editar, mesclar i masteritzar so, per a optimitzar i millorar gravacions de veu, per a reduir soroll i sons no desitjats, per a resumir i editar vídeos obtenint talls amb els millors moments i per a optimitzar aquests continguts editats per a la seua utilització en xarxes socials depenent de la plataforma en la qual es desitgen publicar.

L'objectiu és, una vegada estudiades les opcions, obtindre un procediment estàndard a seguir en la pràctica quan es té un contingut audiovisual en cru (sense processar) per a editar-lo i prepararlo de la manera més senzilla i ràpida possible utilitzant aquestes eines existents.

Es planteja, a més, una reflexió adicional sobre l'ús i possibilitats de la Intel·ligència Artificial en aquest sector en particular.

Abstract

In this work, research has been conducted on different applications for audiovisual editing based entirely or partially on Artificial Intelligence, analyzing the strengths and weaknesses of each one and focusing their use towards the rapid and efficient creation of content for social media. Among the applications analyzed are tools for editing, mixing, and mastering sound, optimizing and improving voice recordings, reducing noise and unwanted sounds, summarizing and editing videos to obtain clips with the best moments, and optimizing these edited contents for use on social media depending on the platform on which they are to be published.



The objective is, once the options have been studied, to obtain a standard procedure to follow in practice when having raw (unprocessed) audiovisual content to edit and prepare it in the simplest and quickest way possible using these existing tools.

Additionally, an additional reflection is proposed on the use and possibilities of Artificial Intelligence in this particular sector.

RESUMEN EJECUTIVO

La memoria del **TFG/TFM** del **NOMBRE TÍTULO** debe desarrollar en el texto los siguientes conceptos, debidamente justificados y discutidos, centrados en el ámbito de la **NOMBRE DISCIPLINA**

| CONCEPT (ABET) | CONCEPTO (traducción) | ¿Cumple? (S/N) | ¿Dónde? (páginas) |
|--|---|-------------------|----------------------|
| 1. IDENTIFY: | 1. IDENTIFICAR: | | |
| 1.1. Problem statement and opportunity | 1.1. Planteamiento del problema y oportunidad | S | 1-2 |
| 1.2. Constraints (standards, codes, needs, requirements & specifications) | 1.2. Toma en consideración de los condicionantes (normas técnicas y regulación, necesidades, requisitos y especificaciones) | S | 1-2 |
| 1.3. Setting of goals | 1.3. Establecimiento de objetivos | S | 3 |
| 2. FORMULATE: | 2. FORMULAR: | S | |
| 2.1. Creative solution generation (analysis) | 2.1. Generación de soluciones creativas (análisis) | S | 4 |
| 2.2. Evaluation of multiple solutions and decision-making (synthesis) | 2.2. Evaluación de múltiples soluciones y toma de decisiones (síntesis) | S | 8-16 |
| 3. SOLVE: | 3. RESOLVER: | S | |
| 3.1. Fulfilment of goals | 3.1. Evaluación del cumplimiento de objetivos | S | 22-24 |
| 3.2. Overall impact and significance (contributions and practical recommendations) | 3.2. Evaluación del impacto global y alcance (contribuciones y recomendaciones prácticas) | S | 23 |



Índice

| | | |
|-------------|---|---|
| Capítulo 1. | Introducción..... | 1 |
| 1.1 | Contexto histórico | 1 |
| 1.2 | Situación actual..... | 2 |
| 1.2.1 | Aspectos positivos | 2 |
| 1.2.2 | Aspectos negativos..... | 2 |
| 1.3 | Inteligencia Artificial en el ámbito audiovisual | 2 |
| 1.4 | Planteamiento y descripción del trabajo | 2 |
| Capítulo 2. | Objetivos del trabajo..... | 3 |
| Capítulo 3. | Metodología..... | 4 |
| 3.1 | Propuesta y fase inicial..... | 4 |
| 3.2 | Desarrollo del trabajo | 4 |
| Capítulo 4. | Aplicaciones de IA para audiovisuales | 5 |
| 4.1 | Adobe | 5 |
| 4.1.1 | Adobe Podcast | 5 |
| 4.1.2 | Adobe Express | 5 |
| 4.1.3 | Adobe Rush | 6 |
| 4.1.4 | Adobe Audition | 6 |
| 4.2 | Waves Audio | 6 |
| 4.2.1 | Waves Sibilance..... | 6 |
| 4.2.2 | Waves SoundGrid Systems | 6 |
| 4.2.3 | Waves Clarity Vx..... | 6 |
| 4.3 | Wisecut | 6 |
| 4.4 | Emastered | 7 |
| 4.5 | Izotope Rx..... | 7 |
| 4.6 | Auphonic | 7 |
| 4.7 | Descript | 7 |
| 4.8 | Magisto..... | 7 |
| 4.9 | Lumen5..... | 7 |
| 4.10 | Otras herramientas | 7 |
| Capítulo 5. | Análisis y explicación de las aplicaciones elegidas | 8 |



| | | |
|--------------|--|-------------------------------------|
| 5.1 | Adobe Enhance Speech | 8 |
| 5.2 | Adobe Express | 9 |
| 5.3 | Waves Clarity Vx | 10 |
| 5.4 | Wisecut | 11 |
| 5.5 | Emastered | 13 |
| 5.6 | Auphonic | 14 |
| 5.7 | Descript | 16 |
| Capítulo 6. | Caso práctico: Soundcool y redes sociales | 19 |
| 6.1 | ¿Qué es Soundcool? | 19 |
| 6.1.1 | Usos y proyectos | 19 |
| 6.2 | Planteamiento del caso práctico | 19 |
| Capítulo 7. | Resolución del caso práctico | 21 |
| 7.1 | Procesos básicos: preparación previa del contenido | 21 |
| 7.1.1 | Contenido con voz | 21 |
| 7.1.2 | Contenido musical | 21 |
| 7.2 | Edición de vídeo: resúmenes, clips y momentos de interés | 22 |
| 7.3 | Redacción de textos, posts y noticias | 23 |
| 7.4 | Edición de contenido para redes sociales | 25 |
| 7.5 | Resumen y conclusiones | 25 |
| Capítulo 8. | Ejemplo de aplicación del caso práctico | 27 |
| Capítulo 9. | Conclusiones y propuestas de trabajos futuros | 30 |
| 9.1 | Reflexión y valoración personal | 31 |
| Capítulo 10. | Bibliografía | 32 |
| Capítulo 11. | Apartados de la memoria | Error! Bookmark not defined. |

Capítulo 1. Introducción

1.1 Contexto histórico

Desde la antigüedad existen mitos sobre artefactos artificiales con conciencia propia, pero no fue hasta principios del siglo XX que surgió el concepto de una máquina o robot con la capacidad de pensar de manera independiente. Este concepto estuvo presente en la mente de los científicos y matemáticos de la década de los 50, incluyendo a Alan Turing, que teorizó sobre la posibilidad de que, si los humanos podemos usar la información que tenemos a nuestra disposición para resolver problemas y tomar decisiones, ¿por qué no iba a poder hacerlo una máquina también?

El campo de investigación sobre IA se fundó como tal en el año 1956, cuando Allen Newell, Cliff Shaw y Herbert Simon presentaron el programa “The Logic Theorist”, diseñado para imitar las habilidades de resolución de problemas de los seres humanos. Este programa, que fue presentado en la Universidad de Dartmouth y fue financiado por la Corporación RAND (Research and Development Corporation), es considerado por muchos el primer programa sobre inteligencia artificial y sus asistentes terminaron por convertirse en investigadores líderes en el sector durante las siguientes décadas.

Desafortunadamente, las expectativas y previsiones de conseguir una máquina tan inteligente como un ser humano en un espacio de tiempo de una generación no se cumplieron, los investigadores habían subestimado enormemente las dificultades del proyecto y a pesar de unos años de progreso notable (entre 1957 y 1974) en los cuales la IA creció y con ella los ordenadores y computadores se volvieron más rápidos, más potentes y más asequibles, este progreso se estancó en la década en los años siguientes debido a que esos años iniciales de avances revelaron la realidad de los obstáculos existentes (principalmente la falta de potencia computacional para hacer cualquier cosa relevante). [1]

A partir de la década de los 80 se vivió otro “boom” debido principalmente a dos factores: una expansión de las herramientas algorítmicas y mayor financiación, se popularizó el término “Deep learning” (lo que permitió a los ordenadores aprender usando la experiencia) y se introdujeron los programas “expert system”, que imitaban los procesos de tomas de decisiones de un humano experto. Este nuevo escenario llevó a una mayor inversión y a un amplio uso de estos sistemas expertos, pero los objetivos marcados y las expectativas generadas durante este periodo no se cumplieron y a partir del año 1987 se creó otro “invierno de IAs”, el mercado de máquinas especializadas de Inteligencia Artificial colapsó y los ordenadores de IBM y Apple superaron en potencia a otras máquinas más caras, los sistemas expertos se volvieron demasiado caros de mantener y su aprendizaje y actualización se volvió inviable.

En los 90 todo cambió, y los objetivos y metas más antiguos planteados para las Inteligencias Artificiales por fin se cumplieron debido principalmente a un incremento drástico de la potencia computacional: en 1997 el programa de IBM Deep Blue le ganó una partida de ajedrez al campeón del mundo Gary Kasparov, se implementó un software de reconocimiento de voz desarrollado por Dragon Systems en Windows y el robot Kismet fue capaz de reconocer y mostrar emociones humanas. [2]

Todo esto nos lleva al siglo XXI, en el cual la Inteligencia Artificial se ha acercado más que nunca a los usuarios promedio. El auge del Big Data (acceso a enormes cantidades de información y datos), los ordenadores más potentes y más baratos que nunca y el interés por estas tecnologías han generado un nuevo auge y ascenso de la Inteligencia Artificial, hasta que en el año 2020 surgieron las plataformas y herramientas de uso masivo que son conocidas y usadas por muchísimos usuarios en la actualidad, modelos como GPT y OpenAI han permitido a cualquier persona con acceso a Internet y un dispositivo utilizar estas herramientas para prácticamente cualquier uso (generación de texto, de imágenes y vídeos, programación y cualquier duda o pregunta que uno pueda plantearse). [3]

1.2 Situación actual

En la actualidad, se ha multiplicado exponencialmente el uso de modelos de inteligencia artificial, principalmente de chatbots basados en LLMs (Large Language Model), pero también generadores de imágenes, música... Esto ha propiciado una serie de consecuencias que han influido en muchos aspectos de la manera en la que nos relacionamos y trabajamos con la información y con Internet, llegando incluso a ser parte primordial en la forma de trabajar de muchas personas. [4]

1.2.1 Aspectos positivos

Los principales aspectos positivos de este “boom” en el uso, prevalencia y distribución de diferentes modelos de IA son, evidentemente, la facilidad de acceder a ellos, su gratuidad en muchos casos, su potencia, y las enormes posibilidades de uso: desde buscar información, apoyo para redactar documentos, traducción, capacidad de que asistan en el trabajo diario... más personas que nunca utilizan diferentes iteraciones de la IA en su vida cotidiana y con asiduidad, lo cual, sin duda, es un avance enorme respecto a décadas anteriores.

1.2.2 Aspectos negativos

También hay que mencionar que no todo lo que rodea a la IA y sus aplicaciones es bien recibido y está libre de controversia. Las altas capacidades computacionales y el uso masivo de estos modelos tienen un impacto medioambiental elevado, además, la mayor crítica hacia estos modelos es seguramente su utilización para fines artísticos como generación de dibujos o animaciones de manera “artificial” y las implicaciones morales que eso puede tener con respecto a los artistas humanos reales y la “invasión” de su sector de trabajo. En esa línea, también se ha planteado la posibilidad de que estas herramientas puedan sustituir en no demasiado tiempo a trabajadores humanos en sectores como la programación o generación de código, ya que estos modelos aprenden de manera muy rápida y los avances vistos en muy poco tiempo han sido enormes. Todo esto ha generado una situación en la cual hay un debate ético y moral respecto a si se debe impulsar el avance y propagación de la Inteligencia Artificial sin pensar en las consecuencias o, si por el contrario o hay que tener cuidado porque los efectos pueden ser difíciles de predecir y de subsanar.

1.3 Inteligencia Artificial en el ámbito audiovisual

Como se ha mencionado en el resumen, este trabajo trata sobre la Inteligencia Artificial aplicada a usos audiovisuales, y un análisis de las herramientas existentes. Estas herramientas sirven para editar audio, vídeo, para subtítular automáticamente, para reducir ruido, adaptar contenido a redes sociales, generar resúmenes de manera automática y en general reducir la carga de trabajo y facilitar los procesos necesarios a los que normalmente se somete esta clase de contenido.

Existen numerosas aplicaciones y herramientas, desde gratuitas hasta de pago, que cumplen este cometido, al igual que en otros sectores y campos de aplicación, lo audiovisual no se ha quedado atrás a la hora de generar herramientas que utilicen la potencia de la IA, sobre todo en esta época actual donde la creación de contenido digital es tan popular y masiva, y se consume por casi todo el mundo de manera diaria.

1.4 Planteamiento y descripción del trabajo

Este trabajo se basa en investigar y analizar las herramientas de Inteligencia Artificial que se utilizan para la creación, edición y generación de contenido audiovisual, se probarán unas cuantas de ellas, se discutirán sus fortalezas y debilidades, se explicará brevemente cómo se utilizan y cómo funcionan y finalmente se planteará un caso práctico enfocado a la plataforma Soundcool de manera que se encuentre una forma lo más rápida y óptima posible de editar el contenido y prepararlo para su distribución en las redes sociales, haciendo énfasis en el uso de Soundcool para salud y educación.



Capítulo 2. Objetivos del trabajo

En este capítulo se tratarán y explicarán los objetivos del trabajo, que principalmente son dos:

- Realizar una investigación sobre las herramientas basadas en Inteligencia Artificial enfocadas al uso audiovisual existentes, elegir las más interesantes para su análisis y plantear una serie de conclusiones tras haberlas utilizado y probado, teniendo en cuenta sus ventajas y desventajas y sus fortalezas y debilidades.
- Plantear un caso práctico cuya que se basa en, teniendo un vídeo en crudo (sin editar de ninguna manera), encontrar la forma más rápida y sencilla de editarlo para utilizarlo en diferentes redes sociales con una calidad razonablemente alta.

El resultado final será una explicación sobre cómo funcionan estas plataformas, qué ventajas e inconvenientes tienen, su posible impacto en la industria, y una breve guía de cómo editar un contenido concreto dependiendo de la red social en la que se desee publicar utilizando estas herramientas y programas. Todo esto enfocado a la plataforma Soundcool, que más adelante se explicará lo que es y cómo funciona.

Como complemento a los dos objetivos principales, se intentará hacer un análisis más personal del estado actual de las IA, en concreto en este sector audiovisual, y las potenciales connotaciones tanto positivas como negativas de la aparición de estas herramientas. A su vez, se expondrán riesgos que su uso podría conllevar, como la cesión de datos.

Finalmente se sacarán una serie de conclusiones generales respecto a los conocimientos obtenidos tras la realización de este trabajo.

Capítulo 3. Metodología

En este capítulo se procede a explicar cuál ha sido la metodología para la realización de este trabajo, desde la propuesta inicial, pasando por las adaptaciones y cambios realizados posteriormente para, finalmente, explicar con detalle el proceso utilizado para cumplir los objetivos propuestos anteriormente.

3.1 Propuesta y fase inicial

Este trabajo fue sugerencia de mi tutor, Jorge Sastre Martínez, cuya propuesta inicial fue realizar un análisis de las herramientas existentes basadas en IA para editar audio y vídeo, probarlas y de alguna manera realizar una guía de uso exponiendo los aspectos positivos y negativos de cada una. Tras este primer acercamiento, se planteó la posibilidad de centrar el análisis en herramientas que redujesen el ruido en material grabado con medios habituales (como por ejemplo un móvil) y que también priorizaran la calidad de la voz en dicha grabación.

Tras varias conversaciones y un intercambio de ideas sobre el tema, se llegó a la conclusión de que lo más interesante sería realizar dicho análisis, pero a su vez aplicarlo a una finalidad práctica y que tuviese una utilidad, fue entonces cuando se decidió aprovechar la plataforma Soundcool y el material existente en ella para enfocar el caso práctico a optimizar un proceso que permita crear y editar contenido para su uso y publicación en redes sociales.

3.2 Desarrollo del trabajo

Una vez establecido cuál iba a ser el tema del trabajo y cuáles iban a ser los objetivos, el siguiente paso fue empezar el trabajo en sí. Dadas las características de este, para empezar, se realizó una breve investigación sobre qué herramientas existían que se adecuaban para los intereses planteados, se probaron dichas aplicaciones y se recopiló una lista de aquellas que resultaban más interesantes y que serían probadas con más exhaustividad para el cumplimiento de los objetivos.

Tras tener claro cuáles eran los programas más interesantes, el siguiente paso fue probarlos con mayor detenimiento, analizando sus ventajas y desventajas, su curva de dificultad para su uso, y compilar un breve manual de uso para cada uno de ellos, de esta manera se pudo desarrollar el apartado del trabajo dedicado a tratar en sí estas herramientas y que puede servir como material interesante ya por sí solo.

Posteriormente, se trató el segundo objetivo: resolver el caso práctico planteado. Se hicieron diversas pruebas con material de Soundcool con el objetivo de reducir el ruido existente en los vídeos, generar resúmenes de los momentos más interesantes, subtítularlos, generar textos que trataran sobre dichos vídeos, optimizar la voz y otras cosas para mejorar el contenido en crudo. Todo esto fue enfocado a tratar este contenido de las maneras necesarias para su publicación en las redes sociales más interesantes, haciendo una guía del proceso seguido para cada una de ellas y, finalmente, llegando a un proceso estandarizado para obtener los resultados deseados.

Como complemento a los objetivos, se realizó una explicación de la plataforma Soundcool, sus usos, objetivos y puntos de interés.

Finalmente, se plantearon una serie de conclusiones obtenidas tras el aprendizaje y proceso realizados durante la realización de este trabajo, así como una serie de posibles aplicaciones y propuestas para desarrollos futuros.

Capítulo 4. Aplicaciones de IA para audiovisuales

En este capítulo se realizará un breve análisis y descripción de las herramientas basadas en Inteligencia Artificial para uso audiovisual. Se tratarán de manera general para posteriormente analizarlas en mayor detalle en el capítulo siguiente.

4.1 Adobe

Adobe es la primera plataforma que surge como opción, con diversas herramientas que cumplen los requisitos deseados, fue la primera analizada y testeada.

4.1.1 Adobe Podcast

Adobe Podcast es una herramienta existente dentro de la suite de Adobe que utiliza IA para usos enfocados a voz. Dispone de tres variables: Studio, Enhance Speech y Mic Check. De estas tres opciones, solo nos interesa Enhance Speech para este trabajo, pero se realizará una breve descripción de todas porque son igual de interesantes dentro del contexto en el que se trabaja.

Es una plataforma que funciona con una suscripción mensual de pago de 12 euros al mes (con una prueba gratuita de 30 días) pero que también dispone de una versión gratuita con características reducidas, principalmente referido a los límites de uso diario, el tamaño de los archivos y el soporte para vídeo además de audio.

Podcast se usa directamente en el navegador, sin necesidad de descargar e instalar nada, y tiene una interfaz muy fácil de usar, de entender y de manejar, lo cual es una de sus mayores ventajas.

4.1.1.1 Studio

Studio es una herramienta en fase de beta que está enfocada a la grabación de voz, con énfasis en la realización de podcasts. Permite crear plantillas gráficas, generar texto de pistas de voz, introducir música y otras pistas preexistentes de audio. Se trabaja con diversas pistas que se pueden manejar y editar y aunque para los objetivos planteados en este trabajo no se va a usar, es sin duda una aplicación potente y fácil de usar.

4.1.1.2 Enhance Speech

De las tres variables de Podcast esta es la que se utilizará para los objetivos planteados. Esta herramienta permite la subida de audio (y también vídeo en la versión premium) con el objetivo de reducir el ruido ambiente y mejorar la calidad de la voz existente en dicho archivo. Su uso es muy sencillo y rápido, es intuitivo y no dispone de muchas opciones, y se explicará en detalle en el siguiente capítulo ya que es una de las herramientas usadas.

4.1.1.3 Mic Check

Esta herramienta se utiliza para ajustar un micrófono de manera óptima para el uso deseado, mediante el uso de la IA se aconseja al usuario sobre la mejor configuración para el micrófono utilizado, dispone de tres parámetros: ganancia, ruido de fondo y eco, que se presentan con una slider ajustable para cada uno.

4.1.2 Adobe Express

Adobe Express es una herramienta gratuita de Adobe que utiliza IA para, principalmente, crear y compartir contenido gráfico. Tiene un amplio abanico de opciones y posibilidades y plantillas, pero lo más interesante para este trabajo es que permite ajustar y editar contenido para redes sociales, además de ajustar y convertir vídeo y audio. Dispone de planes de pago, pero se puede utilizar (con ciertas limitaciones) de manera gratuita. Se detallarán sus características más adelante.

4.1.3 Adobe Rush

Rush es una herramienta dentro de Adobe Premiere que sirve para la edición de vídeo. Es enteramente de pago (aunque tiene una versión de prueba junto con la suite completa de 7 días) y aunque no se basa al 100% en IA, dispone de características que utilizan la Inteligencia Artificial que son bastante interesantes y se usarán en este trabajo. Entre ellas destacan el ajuste automático de los niveles de audio para balancear, por ejemplo, las pistas de música y voz, el reencuadre automático a las proporciones de aspecto deseadas y el reconocimiento facial y de objetos.

4.1.4 Adobe Audition

Audition es la principal herramienta de Adobe para editar audio, se trata de una DAW (Digital Audio Workstation). Al igual que muchas plataformas de Adobe, es enteramente de pago, pero dispone de una prueba gratuita de 7 días.

Es una plataforma completísima, en ella se puede grabar, editar, mezclar y restaurar audio, dispone de una cantidad enorme de características y es una herramienta muy potente de cara al audio, por esto es ampliamente utilizada por profesionales del sector.

Utiliza Inteligencia Artificial para cosas tales como reducción de ruido y sonidos no deseados, ajustar niveles de audio, ajuste de frecuencias y restauración. Esto agiliza muchísimo el trabajo en comparación a la forma manual tradicional de tratar el audio.

4.2 Waves Audio

Waves Audio es una empresa dedicada a la creación y distribución de tecnologías de procesamiento digital de audio, principalmente plugins, con un amplio catálogo de aplicaciones que utilizan Inteligencia Artificial. Las más interesantes para el trabajo y los objetivos planteados son las siguientes.

4.2.1 Waves Sibilance

Plugin dedicado a detectar y corregir ruidos silbantes en pistas de voz.

4.2.2 Waves SoundGrid Systems

Se trata de una serie de sistemas que procesan audio en tiempo real utilizando algoritmos de IA y la nube.

4.2.3 Waves Clarity Vx

El más interesante para los objetivos planteados, Clarity es una herramienta que reduce el ruido en grabaciones de voz mediante el uso de IA. Identifica y separa la voz de ruidos no deseados y es muy potente mejorando la claridad de la voz. Se analizará en detalle en el siguiente capítulo.

Al ser un plugin, como los demás productos de Waves, se necesita un DAW para utilizarlo. Esto hace que el proceso sea más complicado y laborioso y se aleja un poco de los objetivos de rapidez y simplicidad planteados, aun así, se probará y se valorará.

4.3 Wisecut

Wisecut es una herramienta de navegador y uso gratuito (aunque dispone de planes de pago que amplían sus capacidades) que utiliza la IA para editar vídeo. Su funcionamiento es muy simple y su interfaz muy intuitiva, sus principales características son la detección de fragmentos más interesantes, generación de subtítulos y traducción, eliminación de silencios y generación de música inteligente.

Para ser gratuita, es una herramienta bastante interesante y potente que ofrece unos resultados muy buenos, se utilizará para el proyecto y se dará una explicación detallada en los próximos capítulos.

4.4 Emastered

Herramienta online dedicada a la masterización de pistas de audio, uso gratuito y con diferentes opciones para configuración. Permite subir una referencia junto a la pista que se quiere masterizar para que el resultado sea similar a dicha referencia.

Solo la versión de pago permite descargar los resultados.

4.5 Izotope Rx

Izotope Rx es un toolkit que se utiliza principalmente para limpiar y reparar audio, utiliza machine learning para resolver problemas básicos en grabaciones de sonido, música y post producción. Es una herramienta de pago y requiere de una DAW para ser utilizada.

Su uso es un poco más complejo y dispone de una amplia cantidad de opciones.

4.6 Auphonic

Otra herramienta de navegador, de uso gratuito y con diversas opciones de uso enfocadas a limpiar la voz de una grabación. Tras procesar un vídeo, genera dos archivos: el vídeo procesado y el audio por separado. Las aplicaciones que ofrece son interesantes y los resultados aceptables.

4.7 Descript

Descript es una herramienta utilizada para editar audio y vídeo, su principal característica es la capacidad para transcribir el contenido de los archivos utilizados, esto permite editar de manera directa el texto transcrito y a su vez el audio/vídeo. Posee también la capacidad de eliminar palabras de relleno, está bastante enfocada a la creación de contenido, sobre todo podcasts.

4.8 Magisto

Herramienta propiedad de Vimeo que permite crear y editar vídeos usando Inteligencia Artificial. Posee edición automática, que analiza el contenido e identifica los mejores momentos, crea transiciones y añade efectos visuales.

4.9 Lumen5

Software de edición de vídeo que permite convertir texto a vídeo, optimizar contenido para redes sociales, generar animaciones y efectos. Es especialmente fácil de usar y tiene un enfoque atractivo para el marketing y la educación.

4.10 Otras herramientas

Hay infinidad de herramientas que utilizan inteligencia artificial para edición y creación de contenido audiovisual, esto ha sido un breve resumen de algunas de ellas que, debido a diversos factores como su accesibilidad, facilidad de uso, gratuidad o no y aplicaciones para este proyecto, se han considerado las más interesantes y con mejores perspectivas de uso, pero a continuación se mencionan algunas otras herramientas testeadas interesantes según su uso.

Para vídeo: Runway ML, Pikazo, InVideo, Kapwing, Kamua, Movavi, etc.

Para audio: Anchor, Headliner, LALAL.AI, Magenta Studio, Waves Online Mastering, Masterchannel AI, etc.

Capítulo 5. Análisis y explicación de las aplicaciones elegidas

Después de realizar la investigación sobre una cantidad relativamente alta de aplicaciones audiovisuales que utilicen IA, el siguiente paso era filtrarlas y elegir las más interesantes para probarlas en más detalle y detallarlas en este apartado. Los criterios utilizados para elegir las herramientas que sí iban a ser utilizadas y propuestas como solución al objetivo del trabajo han sido, principalmente, los siguientes:

- Dificultad de uso: se han priorizado herramientas fáciles y rápidas de aprender a usar ya que la intención es realizar un proceso ágil y rápido para generar contenido enfocado a redes sociales.
- Accesibilidad: al existir tantas herramientas, hay una variedad muy grande de formas de usarlas y acceder a ellas. Para este trabajo lo más interesante ha sido centrarse en aplicaciones que no tuviesen mucha complejidad para usar: que fuesen gratis (o con pruebas gratuitas) que no requiriesen instalación o descargas, que dentro de una misma hubiese diversidad de funcionalidades y, en general, que proporcionasen agilidad y utilidad.
- Calidad: a pesar de buscar un proceso rápido y sencillo, se le ha dado prioridad a aplicaciones que diesen resultados con una calidad buena y aceptable.
- Aplicación al objetivo propuesto: muchas aplicaciones resultaban útiles en ámbitos que, si bien son interesantes, no ayudaban demasiado a resolver el caso práctico planteado, por lo tanto, las aplicaciones detalladas a continuación sí resultan útiles para este objetivo.

Las herramientas elegidas son: Adobe Enhance Speech, Adobe Express, Waves Clarity Vx, Wisecut, Emastered, Auphonic y Descript.

5.1 Adobe Enhance Speech

Esta herramienta, como se ha mencionado anteriormente, forma parte de la plataforma Adobe Podcast y resulta muy útil para los objetivos planteados. A continuación, se detalla su uso y los resultados que se obtienen.

Es una aplicación enteramente de navegador [5], se puede usar gratuitamente, pero tiene una prueba de 7 días que amplía sus funciones. Sirve para mejorar la calidad de la voz en una grabación de audio o vídeo, eliminando el ruido de fondo y filtrando las frecuencias que no son relevantes a la voz humana.

Funciona de manera muy sencilla: se sube un archivo (de audio o vídeo) y una vez subido hay un interruptor para activar y desactivar el procesamiento de la voz y una barra deslizante que permite ajustar la intensidad del efecto en función de su porcentaje (al 10% el efecto es menos potente y a medida que se incrementa, se nota más).

En la Ilustración 1 se puede ver la interfaz de trabajo de Enhance Speech, con la selección de efecto y las opciones de subir y descargar los vídeos.

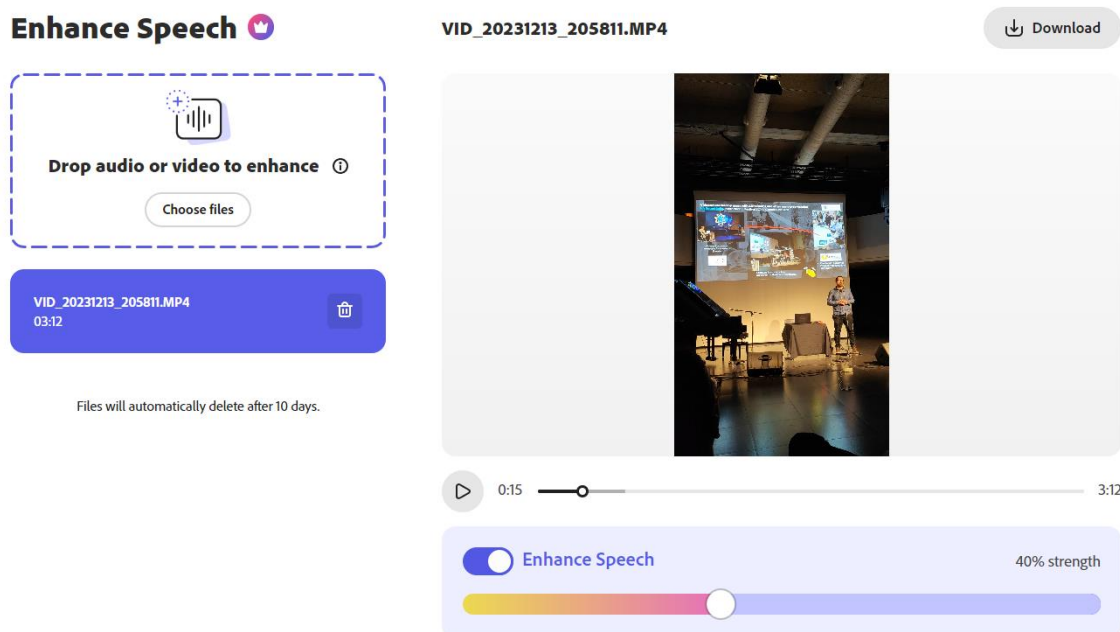


Ilustración 1. Adobe Enhance Speech

Los resultados obtenidos son bastante buenos y debido a su simplicidad es muy fácil encontrar el punto ideal para cada archivo. Funcionando al 100% se nota un audio un poco artificial, pero en el margen entre el 50% y el 80% la verdad es que funciona bastante bien y se obtiene un archivo donde la voz se escucha perfectamente sin ese efecto tan artificial que se produce cuando la herramienta funciona con toda la potencia.

Una vez procesado el archivo, se puede descargar y tiene un tamaño muy similar al archivo original sin procesar.

Las ventajas de esta herramienta son, principalmente, su facilidad y rapidez de uso. No dispone de muchas opciones y si lo que se busca es eliminar ruido de fondo en una grabación en la que haya voz de manera fiable y rápida cumple con esa función. Además, el hecho de que se en navegador y no haya que descargar ni instalar nada, hacen que sea una herramienta muy apropiada para lo que se busca en este trabajo.

Las desventajas observadas son, que quizás la calidad de los resultados no es sobresaliente, que la versión gratuita no permite procesar vídeo y que, si se busca algo sobre lo que tener más control, Enhance Speech solo dispone de la barra deslizante para ajustar el efecto, no tiene ninguna otra opción de ajuste.

Por lo que se ha podido saber, la aplicación no utiliza tus datos para entrenar ninguna IA, lo cual de cara a la privacidad es un punto a favor.

En definitiva, es una herramienta que quizás no es suficiente para un trabajo profesional, detallado y específico, pero para obtener resultados aceptables de manera rápida sin muchos conocimientos previos y generar una cadena de edición de contenido para redes sociales, es, sin duda, muy útil.

5.2 Adobe Express

Express es una herramienta de Adobe que utiliza la Inteligencia Artificial para generar contenido en redes sociales, editar vídeo, fotos, crear documentos (como tarjetas de visita), marketing y numerosas cosas más.

Se utiliza desde el navegador [6] o desde su aplicación para móviles y se puede usar de manera gratuita. Es bastante versátil y permite crear contenido específico para cada red social, lo cual hace que sea una herramienta muy interesante para resolver los objetivos de este trabajo. Se le ha

dado prioridad frente a otros programas de edición ya que, además de su enfoque claro a RRSS, es muy fácil y rápida de usar, a diferencia de otros programas más profesionales y enfocados a un uso más especializado.

En la Ilustración 2 se aprecia la pantalla de trabajo de Adobe Express, con las opciones para preparar contenido para diferentes redes sociales.

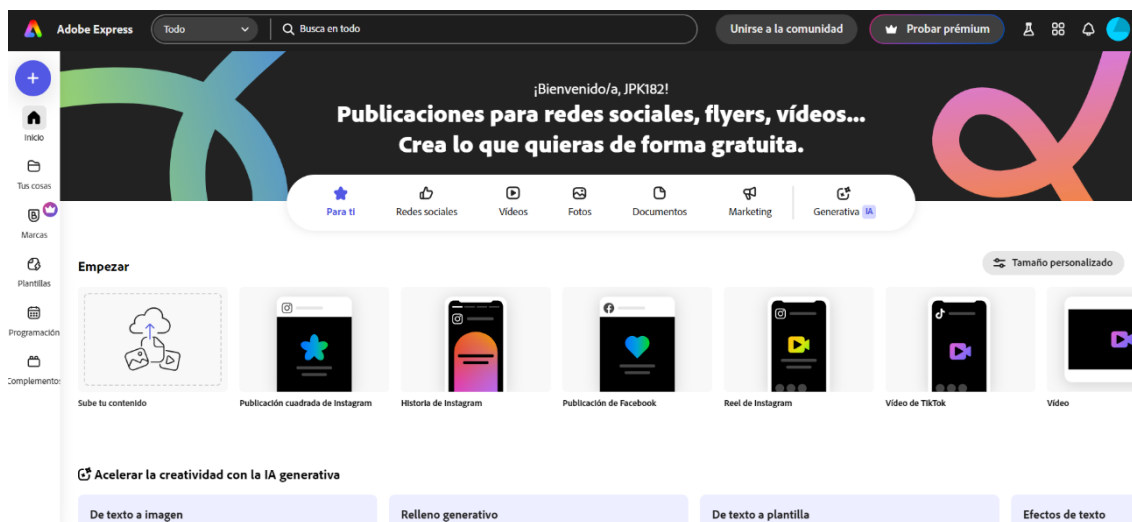


Ilustración 2. Adobe Express

De las herramientas probadas, es la más útil y potente para trabajar con redes sociales.

5.3 Waves Clarity Vx

Clarity Vx es un plugin que forma parte de la suite de Waves Audio, enfocado a reducir el ruido y resaltar la voz en grabaciones de audio. Su funcionamiento es parecido a Adobe Enhance Speech, pero se trata de una aplicación más profesional, con más modos de funcionamiento y opciones, y con una mayor capacidad de controlar los resultados obtenidos.

Es una herramienta de pago (aunque dispone de una prueba gratuita de 7 días), y solo se puede utilizar mediante una DAW (estación de trabajo de audio digital). Para este trabajo se ha probado utilizando Studio One 6 (DAW de pago con 30 días de prueba gratuita).

Su funcionamiento es simple, pero requiere ciertos conocimientos y un proceso previo de configuración, al tratarse de un plugin, hay que vincular la suite de Waves con la estación de trabajo y, una vez estando todo configurado, se puede añadir un archivo de vídeo o audio en una pista dentro de la estación de trabajo y aplicar el plugin sobre la propia pista. Para probarlo, se ha utilizado un vídeo de una grabación de una conferencia, en la cual hay voz y ruido de fondo, la propia estación facilita separar el audio en una pista de trabajo independiente sobre la que aplicar el plugin.

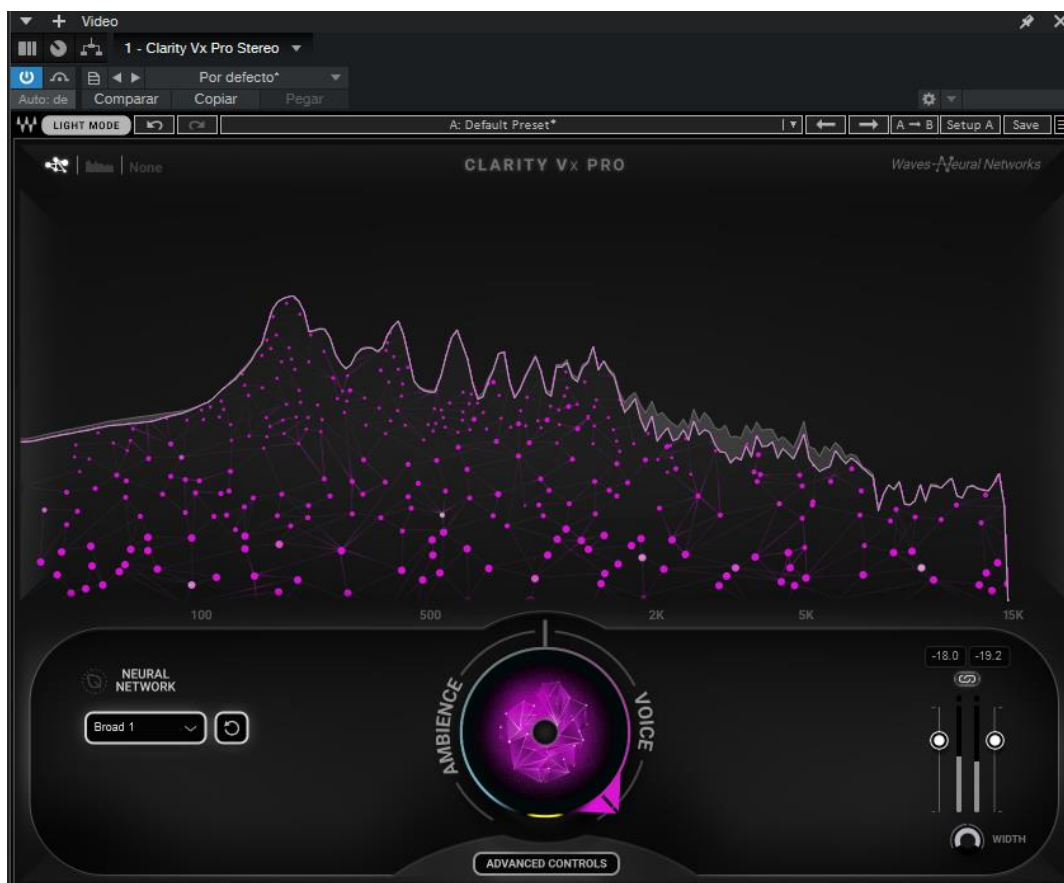


Ilustración 3. Clarity Vx funcionando

Una vez se aplica el plugin sobre la pista deseada se puede ver su funcionamiento en tiempo real, como se aprecia en la Ilustración 3, presentada anteriormente. Los controles son más completos que los de Enhance Speech, podemos observar que hay un potenciómetro con el cual controlar si se potencia el ambiente o la voz, y se aprecia una gráfica de frecuencias en la que se ve la ecualización que está realizando el plugin en tiempo real dependiendo del estado del mando de control. Se puede elegir si funciona en estéreo o mono, y tiene unos presets ya configurados con los cuales se puede elegir un modo de funcionamiento predeterminado.

Los resultados son de una calidad mejor que otras herramientas probadas, se aprecia una mejor calidad de audio resultante y el filtrado es menos artificial y más natural. Al tratarse de un plugin usado en una DAW, se puede configurar de diversas maneras el efecto, utilizando diferentes pistas, envíos auxiliares, niveles e intensidad del efecto en la mezcla final. Una vez obtenido el resultado deseado, se puede exportar el archivo editado con el efecto aplicado.

Las ventajas de esta herramienta son, desde luego, su calidad, profesionalidad y capacidad de controlar manualmente diversos factores y opciones. Está un nivel por encima de otras herramientas probadas (por ejemplo, Enhance Speech).

Por el contrario, sus desventajas son que no es una herramienta gratuita, que requiere software externo para ser utilizada (junto con la correspondiente configuración y vinculación) y que no cumple del todo los objetivos planteados de un proceso rápido pero eficaz. Es, quizás, una herramienta demasiado completa y compleja para lo que se busca en este trabajo.

5.4 Wisecut

Una de las herramientas más interesantes analizadas, se adapta bastante bien a lo que se busca en el trabajo. Es fácil de usar, no requiere instalar nada y funciona de manera razonablemente rápida, proporcionando unos resultados bastante buenos [7].

Se puede utilizar de manera gratuita 60 minutos (de contenido subido) mensuales, con una resolución de 720p y hasta 4 GB de material procesado. Dispone de planes de pago que incrementan estas cifras, pero para lo que se busca, el plan gratuito sobra.

Su funcionamiento es bastante simple, primero se sube el archivo de vídeo que se desea procesar, y una vez la herramienta lo ha subido a su base de datos, los resultados que devuelve son: una transcripción de la voz del vídeo bastante precisa, una serie de hashtags relacionados con la temática del vídeo y varios clips titulados con los momentos más interesantes del contenido.

También se pueden editar los clips, reorganizar las escenas y elegir los fragmentos que se quieren utilizar y tener un control más manual del proceso de creación

Finalmente, se puede exportar el clip o fragmento deseado, y la aplicación incluye la posibilidad de añadir música de fondo generada por IA automáticamente y de añadir subtítulos.

Debido a la simplicidad de su uso, es una herramienta muy útil para resumir vídeos y resaltar momentos destacados, la capacidad de transcribir lo que se dice y de generar resultados con subtítulos desde luego son características muy atractivas y que se ajustan muy bien a la resolución de los objetivos planteados. Además, el funcionamiento de la IA es muy bueno ya que detecta acertadamente de qué trata el contenido para realizar títulos, etiquetas y resúmenes.

Podemos observar en las Ilustraciones 4 y 5 las diferentes interfaces de Wisecut, su página principal y la pantalla de trabajo.

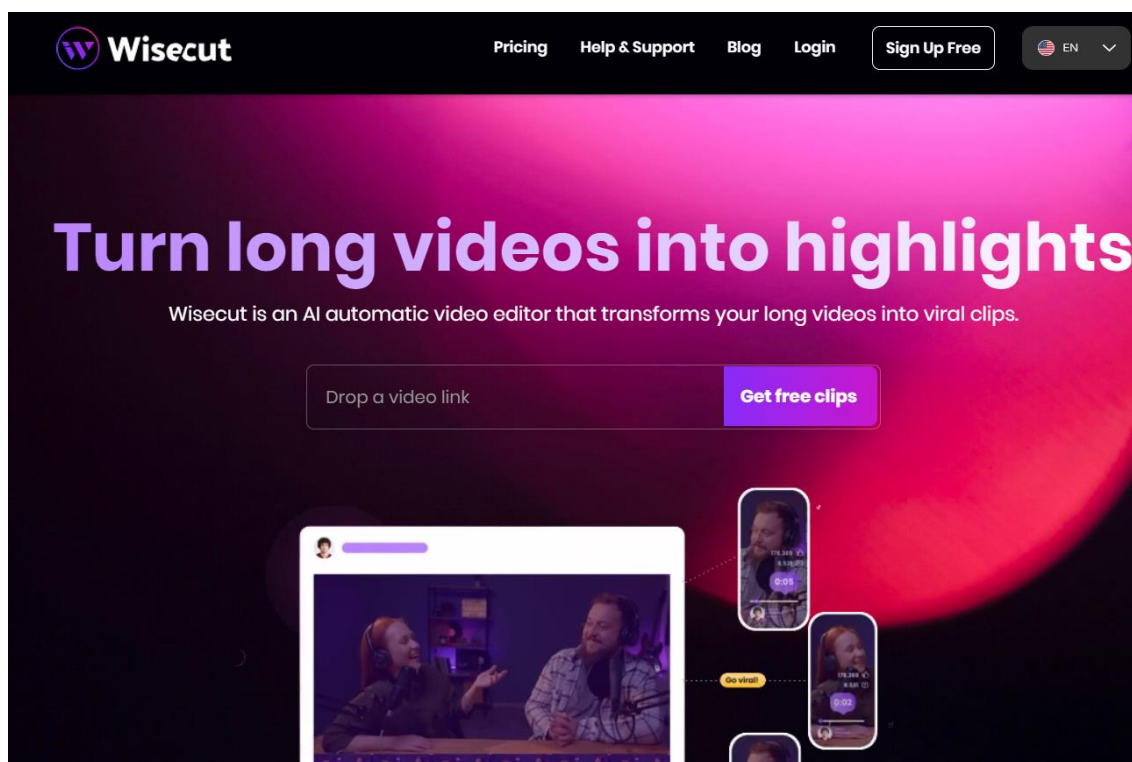


Ilustración 4. Wisecut

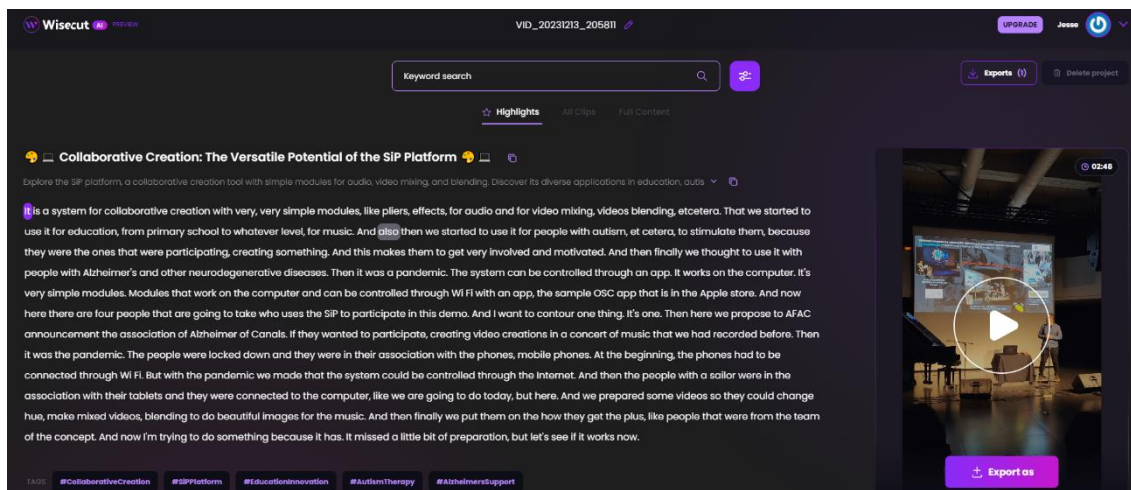


Ilustración 5. Procesado en Wisecut

Respecto al uso de datos, las condiciones de privacidad son las habituales: para hacer marketing, promociones y publicidad a tu cuenta. No queda del todo claro leyendo las condiciones de privacidad si estos datos se utilizan para el entrenamiento de la IA, pero dan a entender que no y que los usos de los datos personales se limitan a los expuestos anteriormente.

En resumen, se trata de una herramienta que ha resultado sorprendente al probarla y que realiza un trabajo bastante útil si se busca simplicidad, rapidez y resultados de calidad.

5.5 Emastered

Emastered permite masterizar archivos de audio usando su sistema de IA online. Se puede configurar de diferentes maneras y es una herramienta bastante simple en esencia, su único uso es realizar masterización.

Tiene utilidad para este trabajo porque en caso de, por ejemplo, querer sustituir el audio grabado con cámara en un archivo de vídeo por una pista de audio con mejor calidad, es una herramienta rápida y fácil para hacer un máster de dicha pista.

En la Ilustración 6 se presentan las opciones de las que dispone Emastered.

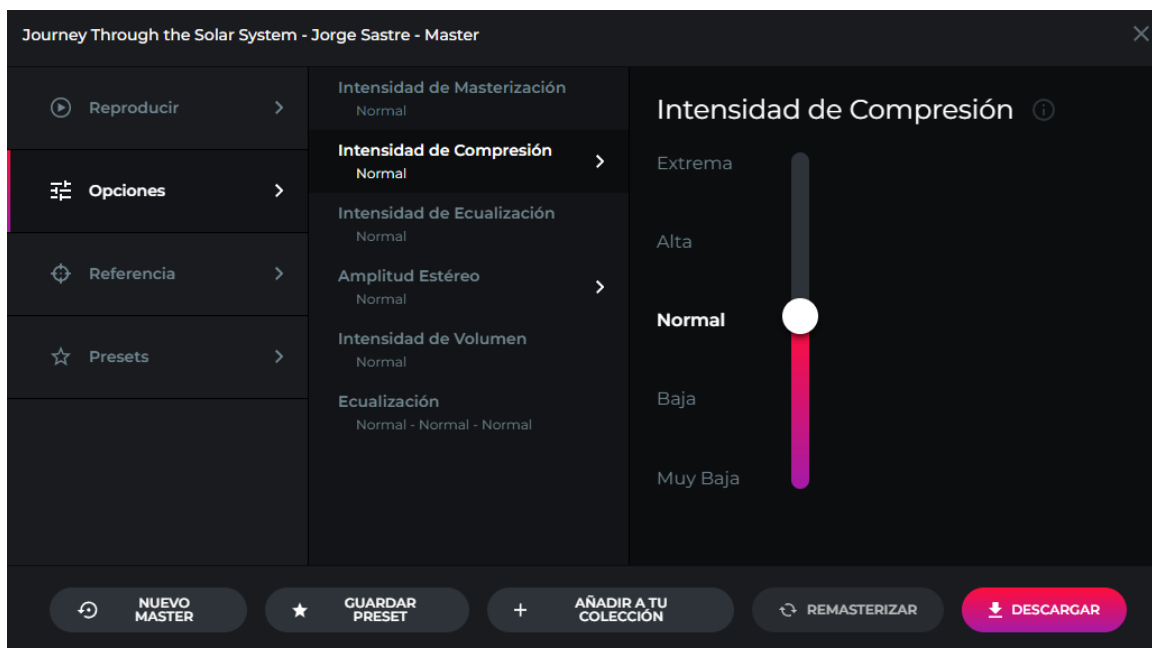


Ilustración 6. Opciones de Emastered

Las diferentes opciones que incluye Emastered son:

- Intensidad del nivel de la masterización: se puede ajustar la intensidad del mastering entre los niveles bajo, medio y alto.
- Intensidad de la compresión y de la ecualización: Emastered permite ajustar el nivel de compresión del sonido y la ecualización de la pista entre muy bajo, bajo, medio, alto y extremo.
- Amplitud estéreo: se puede ajustar la amplitud del espectro estéreo, lo cual ubica los sonidos y objetos espacialmente entre L, Centro y R.
- Intensidad de volumen: se puede ajustar el nivel de volumen del archivo de audio de manera manual.
- Ecualización: se puede modificar la ecualización, ajustando los niveles de las frecuencias bajas, medias y altas por separado.

Todas estas opciones llevan una breve descripción de lo que son y de cómo se utilizan. Además, se pueden crear presets y guardar los parámetros personales de masterización y existe la opción de incluir como referencia una pista cualquiera para que la masterización del archivo procesado sea similar a la pista usada como referencia.

Los resultados se pueden ver de manera gráfica ya que incorpora una comparativa entre la onda sonora del archivo en crudo y del archivo masterizado, como se puede observar en la Ilustración 7.

Emastered no especifica dentro de su política de privacidad si los datos proporcionados se usan para entrenar su IA.

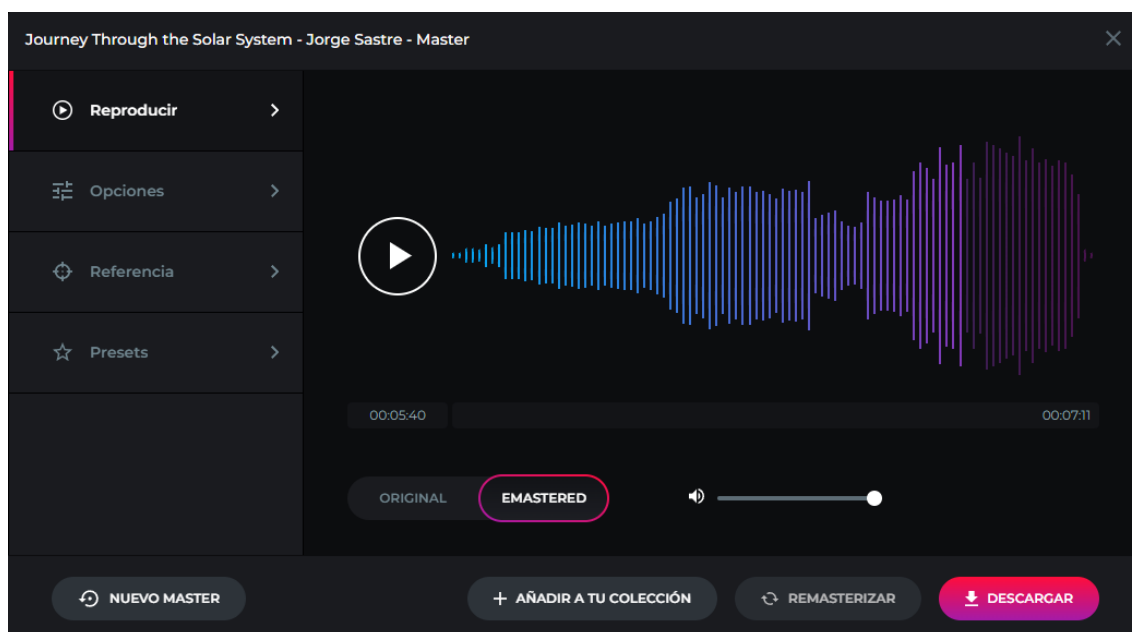


Ilustración 7. Procesado en Emastered

5.6 Auphonic

Auphonic es una aplicación similar a otras analizadas, funciona a través del navegador y se puede utilizar gratuitamente 2 horas cada mes, existiendo planes de pago que amplían ese uso. Dispone de diversas características que van principalmente enfocadas a “limpiar” archivos de audio o vídeo, como, por ejemplo: ajuste inteligente de los niveles de los diferentes elementos de una pista de audio (voz, música...), reducción de ruido, ecualización y filtrado automáticos, eliminar palabras de relleno y transcripción voz a texto, entre otras.

Se puede utilizar para podcasts, para contenido educativo, para audiolibros, creación de contenido... y dispone de una serie de opciones predeterminadas que facilitan el trabajo.

Las opciones de las que dispone Auphonic se pueden apreciar en la Ilustración 8.

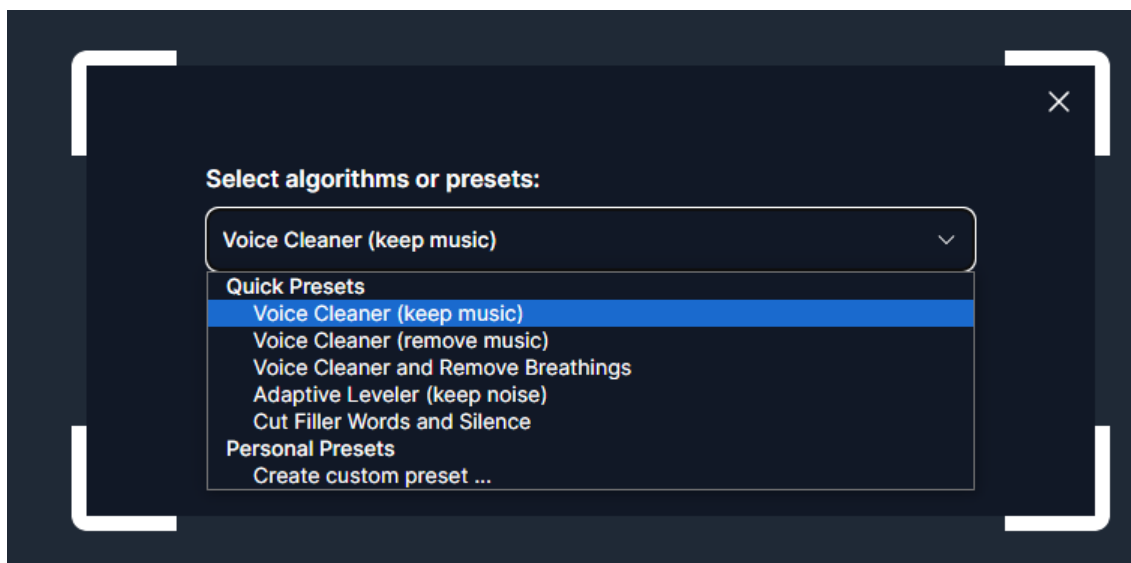


Ilustración 8. Auphonic

Es interesante porque, en caso de procesar un archivo de vídeo, la propia aplicación genera un archivo con el audio por separado, lo cual facilita posibles usos de este. El lado negativo es que la versión gratuita no permite descargar los archivos procesados. Como se aprecia en la Ilustración 9, la aplicación genera una comparativa entre las ondas sonoras del archivo sin procesar y el archivo procesado.

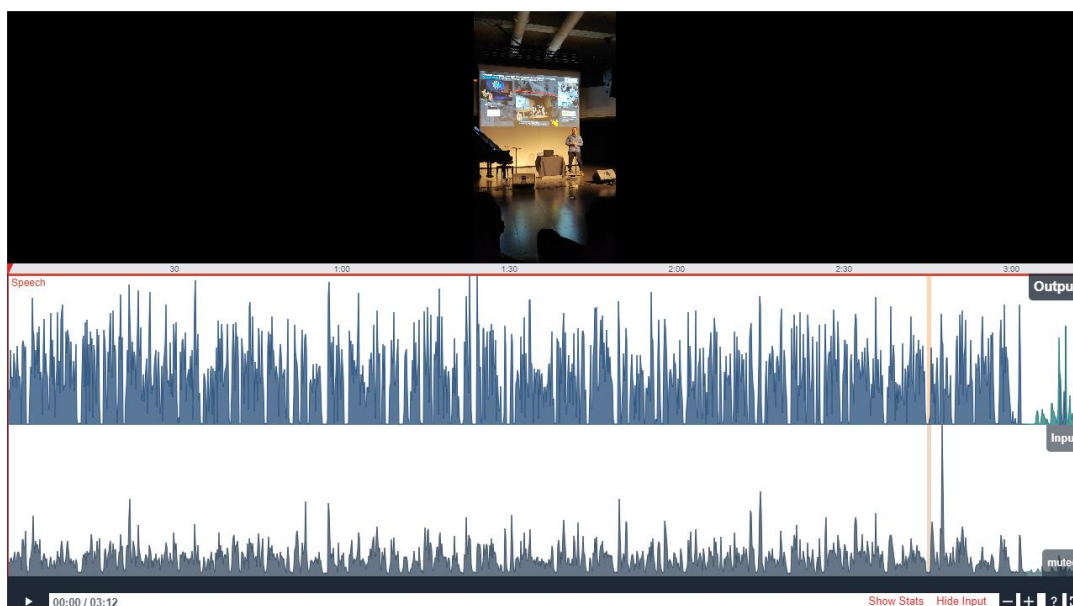


Ilustración 9. Procesado Auphonic

Es, quizás, la herramienta más completa de este tipo y la más polivalente, permitiendo más de una aplicación, todo esto manteniendo un funcionamiento fácil de entender y rápido.

De acuerdo con su política de privacidad, Auphonic no utiliza los datos de usuario proporcionados para entrenar sus modelos de IA, los archivos de audio solo se utilizan para procesarlos y llevar a cabo las tareas requeridas.

5.7 Descript

Descript es una herramienta de edición de vídeo enfocada principalmente en la voz y el texto. Permite subir archivos de vídeo y los transcribe de manera automática, la diferencia principal con otras herramientas probadas es que su principal aplicación es editar y modificar dicha transcripción, eliminar momentos de silencio o palabras de relleno e incluso generar clones de voz con los cuales poder sustituir la voz original del archivo utilizando la IA, de manera que se puede redactar a mano algo que luego será la voz del vídeo.

Es gratuita (aunque tiene planes de pago) y resulta especialmente útil para podcasts, conferencias y contenido educativo. Además, es de uso colaborativo y permite que varias personas trabajen en un mismo proyecto de manera simultánea y se integra fácilmente con otras plataformas de edición.

En la Ilustración 10 se ve la función de transcripción de la que dispone Descript.

Descript no utiliza los datos de usuario para entrenar a su IA, pero se puede optar por activar esta opción y ceder voluntariamente tus datos personales para mejorar la plataforma. La opción viene desactivada por defecto y la elección recae sobre el usuario.

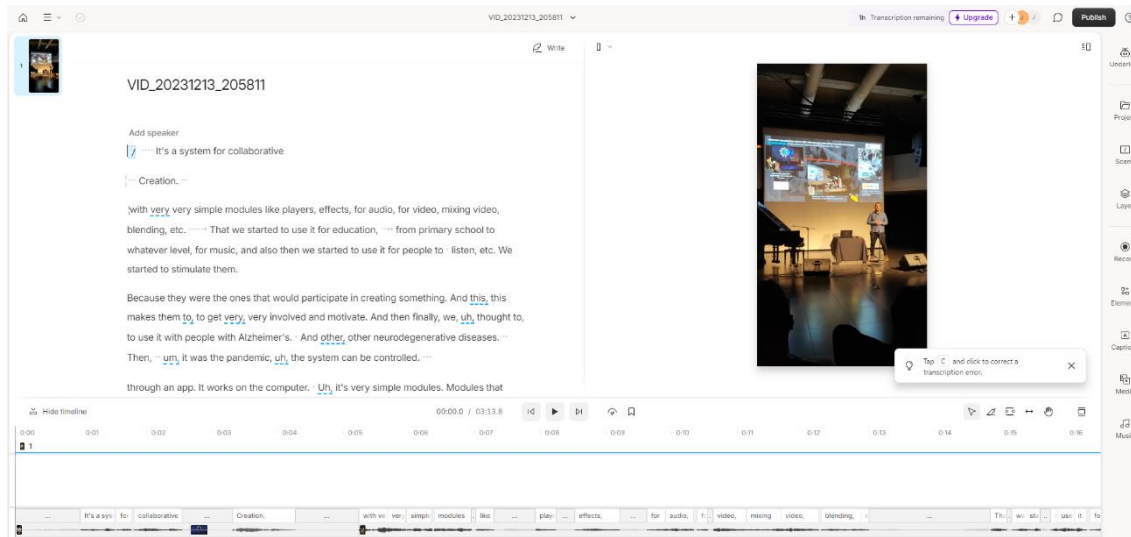


Ilustración 10. Descript

En la Tabla 1 se puede ver una comparativa entre las diferentes aplicaciones mencionadas en este capítulo, en ella se resaltan las características de estas y sus cualidades de Inteligencia Artificial.

| Nombre | Características de IA | Suscripción o pago | Enfoque | Uso de datos de usuario | Modo de uso | Uso en caso práctico |
|----------------|--|--|--|---|--|--|
| Enhance Speech | Utiliza IA para eliminar ruido y mejorar la calidad de la voz en una grabación | Gratis con funcionalidades limitadas. Prueba gratuita de un mes | Enfocada a uso no profesional | No utiliza los datos personales para entrenar su modelo de IA | Uso en navegador | Sí, debido a que para mejorar la calidad de la voz era la herramienta más sencilla |
| Express | Utiliza IA para generar plantillas de publicaciones de redes sociales, ajustar formatos y editar vídeo | Es gratis de usar | Enfoque no profesional, apta para cualquier usuario | No utiliza los datos personales para entrenar su modelo de IA | Se utiliza en navegador, pero se puede descargar una app móvil | Sí, debido a su gratuidad y capacidad de adaptar el contenido a cada red social |
| Clarity Vx | Utiliza IA para ecualizar, detectar y reducir ruido y amplificar las frecuencias de la voz humana | Plugin de pago del paquete de Waves | Enfoque profesional, se necesitan conocimientos básicos de edición de sonido | No utiliza datos personales para entrenar su modelo de IA | Plugin que requiere descarga y una DAW para utilizarse | No, debido a su enfoque profesional, coste y necesidad de utilizar una DAW |
| Wisecut | Utiliza IA para generar clips con momentos relevantes, subtítular y transcribir la voz y recomendar títulos y tags del contenido | Gratis con funcionalidad limitada. Planes de suscripción mensual | Enfoque no profesional, pero dispone de característica avanzadas de edición | No especifica si utiliza los datos de usuario para entrenar su modelo de IA | Uso en navegador | Sí, ya que dispone de características gratuitas |
| Emastered | Utiliza IA para realizar la masterización de pistas de audio | Gratuito con limitaciones | Enfoque no profesional | No usa datos de usuario para entrenar su modelo de IA | Uso en navegador | Sí, ya que es una herramienta útil para ajustar pistas de música |
| Auphonic | Utiliza IA para mejorar la | Gratis con limitaciones | Enfoque semiprofesional | No utiliza los datos | Uso en navegador | No, debido a que es una |



| | | | | | | |
|----------|--|---|--|---|------------------|--|
| | calidad del audio de diferentes maneras (ajuste de volumen, rango dinámico, ecualización...) | | | de usuario para entrenar su modelo de IA | | herramienta algo más compleja de lo que se estaba buscando |
| Descript | Utiliza IA para ajustar y transcribir las palabras de un vídeo. Se pueden realizar correcciones, traducciones y sustituciones automáticas. Genera voz con IA | Gratis con limitaciones, existen planes de pago | Enfoque no profesional, pero es una aplicación medianamente compleja | No utiliza datos de usuario para entrenar su modelo de IA | Uso en navegador | No, debido a que Wisecut era más intuitiva y fácil de usar |

Capítulo 6. Caso práctico: Soundcool y redes sociales

En este capítulo se planteará el caso práctico, se expondrá la manera escogida para resolverlo y se analizarán y valorarán los resultados. Además, se realizará una breve explicación de la plataforma Soundcool y su funcionamiento, siendo la creación de contenido para las redes sociales de esta plataforma el objetivo final buscado.

6.1 ¿Qué es Soundcool?

Soundcool es un sistema de creación audiovisual desarrollado en la Universidad Politécnica de Valencia por el equipo de Jorge Sastre (tutor de este proyecto) y cuya principal característica es que es colaborativo. Utiliza principalmente dispositivos móviles y fue utilizado inicialmente para educación. [8]

Está diseñado para ser accesible a todo el mundo, tanto músicos y artistas experimentados como principiantes sin conocimientos previos, y sus aplicaciones principales son la creación artística, la educación y la salud. [9]

Dispone de una interfaz modular para crear combinaciones de sonido y vídeo y, como se ha mencionado anteriormente, permite la colaboración en tiempo real de diferentes personas a través de diferentes dispositivos.

Su sistema básico es gratuito y ha sido utilizado para una multitud de proyectos internaciones que serán detallados a continuación.

6.1.1 Usos y proyectos

Como se ha comentado, Soundcool se enfoca en proyectos artísticos, educativos y de salud, realizados de manera colaborativa y por profesionales o gente sin formación específica.

Ejemplos de proyectos artísticos desarrollados son la composición de bandas sonoras para cine mudo, creación de efectos sonoros como loops o actuaciones controlando la música mediante Realidad Aumentada con las gafas Hololens.

A nivel educativo, Soundcool tiene muchas posibilidades dadas sus características colaborativas en tiempo real, de manera que, por ejemplo, una clase de alumnos puede participar en un mismo proyecto de manera directa, incluso teniendo cada uno un rol propio. Como ejemplo existente, se realizó un taller consistente en diseñar una mascletà virtual con alumnos de 3º de la ESO, aprendiendo a manejar el entorno Soundcool para componer la propia mascletà.

Finalmente, un uso muy importante de Soundcool es aquel enfocado a la salud, utilizando el sistema para terapias de estimulación cognitiva de cara a enfermedades neurodegenerativas, en particular el Alzheimer. [10]

6.2 Planteamiento del caso práctico

Además del análisis e investigación realizado sobre la Inteligencia Artificial enfocada a usos audiovisuales, este trabajo plantea un caso práctico en el cual se puedan probar dichas aplicaciones de cara a un caso real y poder aplicarlas a la creación y adaptación de contenido en redes sociales.

Se ha elegido Soundcool para este caso práctico ya que la plataforma dispone de numeroso material audiovisual (grabaciones de vídeo, audio, conferencias, conciertos...) y el objetivo marcado es encontrar una forma rápida, fiable y provechosa de editar este contenido disponible utilizando IA para su publicación en las redes sociales de Soundcool, una especie de guía/tutorial para cada caso posible dependiendo de la red social en la que se desee publicar, del tipo de contenido, de si se requiere redactar una publicación o no, y, básicamente, todas las variables que pueden surgir según el tipo de contenido que se busca. Es por esto que se analizarán diferentes



escenarios y se planteará su resolución de manera satisfactoria (por ejemplo, editar un vídeo para subirlo al formato de Shorts de Youtube) utilizando las herramientas previamente analizadas.

Aunque se va a tratar con material de Soundcool, evidentemente estas guías y formas de actuación dependiendo de lo que se desee obtener funcionan con cualquier tipo de contenido, ya que lo importante es explicar cómo editar contenido (usando IA) para redes sociales.

Las redes sociales elegidas para hacer este estudio son las más populares del momento: X (anteriormente Twitter), Instagram, Youtube y Tik Tok. Además, se contempla la posibilidad de redactar posts/noticias/artículos usando estas herramientas.

Se realizará una explicación detallada de la forma que se ha considerado mejor para cada tipo de contenido, ya sea un vídeo corto resumido, un vídeo que requiera subtítulos, una grabación musical, etc.

Capítulo 7. Resolución del caso práctico

En este capítulo se va a plantear la resolución de los objetivos propuestos y del caso práctico desarrollado en el capítulo anterior. Se trata de aplicar las herramientas analizadas de la manera más eficaz para conseguir un flujo de trabajo rápido, sencillo y de calidad que permita adaptar el contenido a las necesidades que se tengan en cada ocasión.

Se detallarán los aspectos positivos y negativos de los procesos realizados y, como se ha mencionado con anterioridad, se priorizará la simplicidad y la rapidez, ya que hay una diferencia metódica entre editar contenido audiovisual para producciones más grandes (proyectos musicales, vídeos largos, conferencias, etc.) que para editar contenido enfocado a redes sociales, situación en la que se busca un fácil acceso al contenido y las herramientas y procesos rápidos enfocados sobre todo al uso de dispositivos móviles.

7.1 Procesos básicos: preparación previa del contenido

Como base de cualquier proceso posterior, se plantea un procesado previo dependiendo de las características del contenido a editar.

7.1.1 *Contenido con voz*

En el caso de que el material a editar contenga o conste de voz hablada (por ejemplo, un clip de una conferencia), se propone empezar el proceso pasando dicho contenido por el programa Adobe Enhance Speech para eliminar ruidos no deseados y limpiar la voz, de manera que se entienda y se escuche adecuadamente como principal prioridad.

La forma de utilizar Enhance Speech es muy sencilla, al acceder a la herramienta en el navegador, se debe subir el archivo de audio o vídeo que se desea editar, una vez ha sido procesado por la plataforma, el usuario dispone de una vista previa para visualizar el contenido, de un botón para activar y desactivar el efecto de mejora de la voz, y de un control deslizante para ajustar la intensidad del efecto (como se aprecia en la Ilustración 1 de este trabajo). Finalmente, tras obtenerse el efecto deseado, se puede descargar el archivo ya procesado.

Con las pruebas realizadas se ha determinado que los resultados utilizando esta herramienta son bastante aceptables y producen un contenido donde la voz se entiende de manera notable, el ruido externo se reduce considerablemente y se obtienen resultados naturales y no demasiado artificiales o extraños.

Como referencia, el margen óptimo aparente para procesar utilizando Enhance Speech es colocar la intensidad del efecto entre el 60% y el 80%, pero evidentemente depende del caso y se pueden probar diferentes combinaciones.

Como alternativa al uso de Enhance Speech, se puede utilizar Waves Clarity Vx si se dispone de una DAW, la suscripción a la suite de Waves y se tienen conocimientos más complejos de edición, esto conllevaría un mayor tiempo en el proceso global pero unos resultados superiores si se utiliza correctamente la aplicación.

7.1.2 *Contenido musical*

En caso de tratarse de un contenido musical o que vaya a utilizar alguna pista con música, el paso primero planteado dentro del proceso sería realizar una masterización de dicha pista utilizando la herramienta Emastered, para equilibrar los niveles, mejorar el rango dinámico y obtener un mejor resultado.

A la hora de utilizar Emastered, no hay que modificar las opciones que vienen por defecto, ya que el resultado obtenido así es bastante aceptable. Sin embargo, si se tiene conocimiento y se desea probar alternativas, la herramienta dispone de diferentes formas de ajustar la masterización de la pista.

También, como parte de un proceso que incluya música, se recomienda utilizar la pista de audio grabada de manera óptima en caso de disponer de ella, y sustituirla por el audio que pueda haber grabado una cámara de vídeo. En el ejemplo práctico de Soundcool este proceso se explica y se realiza.

La utilización de Emastered es muy sencilla y similar a las demás de este trabajo, se sube el archivo a la plataforma, se procesa y una vez finalizado este proceso se puede descargar el archivo ya editado.

Como alternativa o complemento a Emastered, se recomienda el uso de Auphonic, que es una herramienta con un rango más amplio de opciones y de posibilidades más allá de la masterización de las pistas de audio. También se podrían utilizar herramientas más complejas y profesionales como Izotope Rx si el usuario dispone de los conocimientos necesarios.

7.2 Edición de vídeo: resúmenes, clips y momentos de interés

El siguiente paso que abordar es trabajar vídeos de una duración excesiva que puedan generar diferentes clips con momentos interesantes. Para este cometido se utilizará la herramienta Wisecut, se han probado diferentes aplicaciones (como Magisto, Pikazo, Lumen5...) y herramientas que cumplen una función similar, pero la elegida finalmente ha sido ésta, debido a su simplicidad, al hecho de que es gratuita y a los resultados que se pueden generar con ella.

Wisecut va a permitir obtener clips y recortes de un vídeo (destacando los momentos importantes), subtítularlo, añadir música generada por IA si así se desea, y obtener una transcripción, un resumen y una serie de etiquetas del vídeo procesado, lo cual se puede utilizar para generar posts, noticias o el texto que acompañe a la publicación en la red social deseada.

Los clips generados del vídeo resumido se pueden utilizar posteriormente como base para lo que se va a publicar en las redes sociales.

De esta manera, Wisecut formaría parte del proceso de preparación del contenido antes de enfocarlo a una red social concreta. A pesar de existir una gran cantidad de herramientas con un funcionamiento parecido, finalmente se ha decidido contar con Wisecut porque es rápida, fiable, cuenta con diversas opciones de personalización y su plan gratuito deja bastante margen de maniobra para crear contenido.

La utilización de Wisecut es sencilla e intuitiva, se trata de una herramienta de navegador y el primer paso es subir el archivo que se desea procesar a la plataforma, una vez subido, la aplicación genera automáticamente una serie de clips y recortes con momentos destacados y elige uno de ellos como recomendación. Se hace una transcripción del vídeo, y a cada clip se le asigna un título y unos hashtags. Se pueden realizar ediciones como correcciones ortográficas, recortes y edición de vídeo un poco más avanzada (filtros, montajes, etc). Una vez se ha seleccionado el clip que se va a utilizar, se puede exportar el mismo y añadirle música, subtítulos automáticos, marcos y otra serie de efectos y finalmente, cuando se tiene el resultado deseado, la plataforma lo genera en un proceso que tarda aproximadamente unos 15 minutos y permite su descarga con todos los efectos y procesos incluidos.

Existen diversas alternativas a Wisecut ya que hay multitud de editores de vídeo con funcionalidades similares. Se puede utilizar Descript, Magisto, Lumen5... algunas de ellas incluyen también características para enfocar el contenido a redes sociales. Se ha elegido Wisecut porque era la herramienta más sencilla de usar, pero hay un amplio catálogo de posibilidades que se pueden ajustar a diferentes usuarios, procesos y necesidades, siendo Descript la que se ha probado en este trabajo como alternativa.

7.3 Redacción de textos, posts y noticias

Aunque en principio el objetivo y razón del trabajo es analizar herramientas que utilicen IA para editar contenido audiovisual, debido a la finalidad de utilizar este contenido en redes sociales se plantea este uso adicional como complemento: la utilización de la IA para redactar publicaciones que acompañen al contenido audiovisual.

Se han probado diferentes herramientas para este cometido, ya que la cantidad de aplicaciones existentes que se centran en texto son muy numerosas, pero finalmente se ha optado por usar directamente ChatGPT, debido a que es una herramienta fácil de usar y que introduciendo los prompts y comandos concretos permite redactar noticias sobre un tema deseado si se utiliza correctamente.

Como tiene “memoria”, se puede ir editando sobre la marcha el texto generado, realizando correcciones e indicando a ChatGPT cómo modificar cada parte del proceso. En la ilustración 10 se puede ver el primer paso del proceso, donde se le indica a ChatGPT de manera genérica que redacte un texto para una publicación en Instagram.

🎵 ✨ ¡La música está en constante evolución y queremos hablar de ello! 🎧

No te pierdas nuestra Conferencia sobre el Estado Actual de la Música, donde analizaremos cómo la industria musical está cambiando, los géneros emergentes y cómo la tecnología está redefiniendo la creación y distribución musical. 🎧 🚀

📅 Fecha: [Fecha de la conferencia]

🕒 Hora: [Hora del evento]

📍 Lugar: [Ubicación del evento]

🎫 Entradas disponibles en: [Enlace a las entradas]

Ven a escuchar a expertos, artistas y productores compartir sus perspectivas sobre lo que está marcando tendencia y lo que se viene para el futuro de la música. ¡Es una oportunidad única para estar al tanto de lo último en el mundo musical! 🔥 🎧

#Música #Conferencia #EstadoDeLaMúsica #IndustriaMusical #TendenciasMusicales #Música2024 #InnovaciónMusical #EventosMusicales

Puedes ajustar la fecha, hora y detalles de acuerdo con el evento que quieras promocionar. ¿Te gustaría agregar algo más?



Ilustración 10. Petición a ChatGPT

Como se puede apreciar, al pedirle a ChatGPT que redacte un post de Instagram sobre el estado de la música actual, el resultado es una base genérica sobre la que se puede trabajar para ir detallándolo más. La IA ha añadido emojis y los datos sobre el hipotético evento.

El siguiente paso va a ser pedirle a ChatGPT que añada la información concreta (fecha, hora, lugar y la no necesidad de adquirir entradas), el resultado se aprecia en la ilustración 11.

🎧🔗 ¡La música está en constante evolución y queremos hablar de ello! 🎧🎧

No te pierdas nuestra **Conferencia sobre el Estado Actual de la Música**, donde analizaremos cómo la industria musical está cambiando, los géneros emergentes y cómo la tecnología está redefiniendo la creación y distribución musical. 🎧🚀

📅 **Fecha:** 18 de septiembre

🕒 **Hora:** 18:00

📍 **Lugar:** Sala de conferencias de la ETSIT, UPV

🎫 **Entrada libre**, pero recuerda que el aforo es limitado. ¡Llega temprano para asegurar tu lugar! ⌚

Ven a escuchar a expertos, artistas y productores compartir sus perspectivas sobre lo que está marcando tendencia y lo que se viene para el futuro de la música. ¡Es una oportunidad única para estar al tanto de lo último en el mundo musical! 🔥🎧

#Música #Conferencia #EstadoDeLaMúsica #IndustriaMusical #TendenciasMusicales #Música2024
#InnovaciónMusical #EventosMusicales

Ilustración 11. Modificación post ChatGPT

El siguiente paso será añadir datos más concretos sobre la temática de la hipotética conferencia, proporcionar más detalles y que sea menos genérico. Se añade una mención a la Inteligencia Artificial y a un invitado sorpresa, y se puede ver como ChatGPT ajusta cada vez los tags incluidos. En la ilustración 12 se presenta el resultado final.

🎧🔗 ¡La música está en constante evolución y queremos hablar de ello! 🎧🎧

No te pierdas nuestra **Conferencia sobre el Estado Actual de la Música**, donde analizaremos cómo la industria musical está cambiando, los géneros emergentes, el impacto de la **Inteligencia Artificial** en la **producción musical**, y cómo la tecnología está redefiniendo la creación y distribución de música. 🎧🎧🚀

Además, contaremos con la participación de **un invitado sorpresa** que ha revolucionado la forma de componer y está marcando un antes y un después en la industria. ¡No te lo puedes perder! 🎧🎧

📅 **Fecha:** 18 de septiembre

🕒 **Hora:** 18:00

📍 **Lugar:** Sala de conferencias de la ETSIT, UPV

🎫 **Entrada libre**, pero recuerda que el aforo es limitado. ¡Llega temprano para asegurar tu lugar! ⌚

Ven a escuchar a expertos, artistas y productores compartir sus perspectivas sobre lo que está marcando tendencia y lo que se viene para el futuro de la música. ¡Es una oportunidad única para estar al tanto de lo último en el mundo musical! 🔥🎧

#Música #Conferencia #EstadoDeLaMúsica #IndustriaMusical #TendenciasMusicales #Música2024
#InnovaciónMusical #InteligenciaArtificial #ProducciónMusical #InvitadoSorpresa #EventosMusicales

Ilustración 12. Post final

Tras este ejemplo se puede apreciar cómo ChatGPT es una herramienta perfectamente válida para cubrir las necesidades que puedan aparecer a la hora de redactar publicaciones. A pesar de la

existencia de plataformas más específicas, ChatGPT es una opción perfectamente válida y recomendada, permite editar, volver a ediciones anteriores, añadir detalles y, en resumen, utilizar su IA para ajustar una publicación hasta el mínimo detalle.

Un ejemplo curioso e interesante de este uso ocurrió recientemente, cuando el conocido youtuber Spursito utilizó ChatGPT para crear un post y publicarlo en X (anteriormente Twitter) sobre un fichaje falso de un futbolista conocido a nivel mundial, dicho post se hizo viral y mucha gente del sector se lo creyó y lo dio por válido. Se puede apreciar en el vídeo [11] cómo este creador de contenido genera y modifica el post de una manera similar a la que se ha ilustrado en este trabajo.

Como alternativa de uso a ChatGPT para redactar publicaciones, se puede utilizar cualquier bot de texto basado en IA, se puede utilizar Descript o herramientas como Copy.ai, Jasper o Writesonic, todos editores de texto con plantillas disponibles y funciones de IA para ayudar en el proceso.

7.4 Edición de contenido para redes sociales

Finalmente, para editar el contenido y dejarlo listo para su publicación en redes sociales, se ha optado por utilizar Adobe Express. Originalmente, la intención era encontrar una aplicación o herramienta específica diferente para cada red social de interés, planteándose en un inicio la posibilidad de que hubiese herramientas que funcionasen mejor para unas redes sociales que otras, pero, finalmente, tras realizar la investigación y probar las opciones disponibles se valoró que la mejor opción y la más consecuente con este trabajo y sus objetivos era Adobe Express.

Dentro de una sola herramienta se tiene la posibilidad de editar y crear una publicación para todas las redes sociales elegidas como interesantes (Instagram, TikTok, Facebook y Youtube), no hay un modo de funcionamiento específico para X (anteriormente Twitter) pero ésta red no tiene un formato específico y no hay un límite en la duración de los vídeos (aunque existen recomendaciones que realiza la plataforma sobre el formato, tamaño, resolución y otras características) [11], para las demás, las que requieren un formato de tamaño o duración concreto, Adobe Express cuenta con la posibilidad de adaptar el contenido a cada una de ellas.

Express dispone de una gran cantidad de funciones basadas en IA (plantillas, marcos, filtros...) pero en el caso que nos ocupa, como se dispone de un material audiovisual ya grabado y preprocesado, se utilizará la Express subiendo este material a la plataforma y simplemente ajustando su formato (duración, resolución, ratio de aspecto y etc) a la red social deseada, algo que Express realiza automáticamente con su modelo de IA.

Se ha valorado muy positivamente este aspecto de Express, ya que se busca simplicidad, rapidez, agilidad y la posibilidad de trabajar con el contenido desde el menor número de plataformas, así pues, al contrario de lo planteado originalmente, se propone la utilización de Adobe Express como la única herramienta para preparar el material de manera específica a las redes sociales deseadas.

Como alternativa a Express, se puede utilizar Canva (aunque sus funciones premium son de pago), Lumen5, Later o Piktochart. Pero, se ha valorado y finalmente elegido Express para este trabajo porque aglutina muchas de las funcionalidades de otras plataformas en una sola, además de ser capaz de preparar el contenido específicamente para la red social deseada.

7.5 Resumen y guía de trabajo

En este apartado se ha planteado un marco (una guía) fácil de seguir de cara a editar un contenido que ultimadamente será publicado en redes sociales, valorando cada posibilidad encontrada en el proceso y recomendado una serie de herramientas que, en comparación con otras que han sido probadas, han destacado en los aspectos valorados en este trabajo.

La cantidad de herramientas existentes que de una manera u otra cumplen los objetivos es enorme, hay infinidad de plataformas, aplicaciones, páginas y software que utilizan inteligencia artificial

para editar contenido audiovisual, para su publicación en redes, y, en definitiva, para facilitar el trabajo humano y agilizar estos procesos. Ha sido un proceso laborioso probarlas, analizarlas y sacar conclusiones de ellas, pero finalmente se ha intentado mantener un criterio para plantear una genuina recomendación informada y reducir esa cantidad tan grande a un número reducido, accesible y gestionable de herramientas que utilizar en el proceso práctico planteado.

En el Diagrama 1, expuesto a continuación, se ilustra el proceso, paso a paso, recomendado y las aplicaciones sugeridas. Se describe el procedimiento para preparar un archivo de vídeo ya que es el ejemplo más completo y el formato más habitual para publicar en redes. Cabe mencionar que cualquier proceso creativo es enteramente subjetivo y no todos los usuarios tienen por qué valorar igualmente estas recomendaciones, pero como esquema práctico se trata de un resumen útil e informativo, aun así, en los apartados anteriores se proponen alternativas a las aplicaciones recomendadas en esta guía, en caso de que el usuario tenga otras necesidades o sea capaz de utilizar herramientas más complejas y avanzadas.

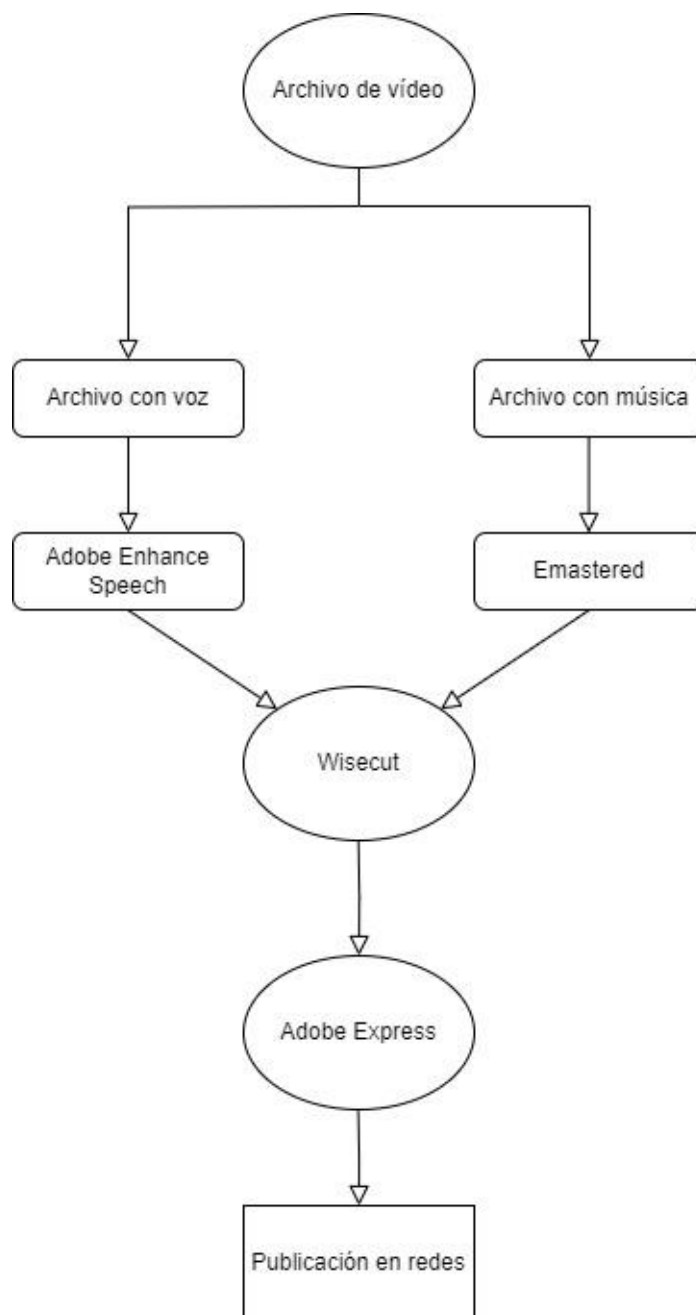


Diagrama 1. Pasos para editar un video

Capítulo 8. Ejemplo de aplicación del caso práctico

Para finalizar el trabajo se van a utilizar los procesos explicados en el capítulo anterior para editar y preparar un archivo específico con las herramientas propuestas, de manera que se demuestre la forma de trabajar a la que se ha llegado como resolución del objetivo que se planteaba. Este proceso se va a explicar a continuación, y durante presentación de este trabajo se hará una demostración visual de los resultados obtenidos y la metodología empleada.

Al final de este capítulo se adjuntarán enlaces de Google Drive que incluyen vídeos que muestran los resultados de los diferentes pasos de este proceso.

El vídeo elegido para editar es un fragmento de tres minutos y doce segundos de una conferencia sobre Soundcool en la cual Jorge Sastre, tutor de este proyecto, explica brevemente el funcionamiento de la plataforma y sus aplicaciones como estímulo y uso para personas con enfermedades neurodegenerativas (en particular Alzheimer), se trata de un vídeo grabado desde el público con un teléfono móvil que tiene algo de ruido ambiente no deseado y la voz se escucha con un volumen relativamente bajo y del que se pueden extraer algunos clips relevantes de una duración reducida para mejorar su formato de cara a una red social.

Como se ha mencionado en el capítulo anterior, el primer paso a realizar tratándose de un archivo en el que se desea priorizar la voz es procesarlo con Adobe Enhance Speech con un nivel del 50% de intensidad, esto dejará un archivo en el que la voz se puede escuchar con mucha más claridad y el ruido exterior se vuelve casi inexistente. El volumen de la voz se ha incrementado, se ha normalizado, y el ruido ha sido reducido, todo esto lo realiza la aplicación automáticamente ya que no dispone de opciones concretas para ajustar estos niveles, solo la intensidad del efecto. En la ilustración 1, en la página 9, se aprecia la pantalla de Enhance Speech procesando el vídeo.

El siguiente paso es utilizar el vídeo procesado con Enhance Speech y trasladarlo a Wisecut, como se ha mencionado antes, esta herramienta genera automáticamente clips con los momentos más relevantes y realiza una transcripción automática. En la ilustración 11 se puede ver el resultado de procesar el vídeo con Wisecut, se ha generado un clip de menor duración, se ha realizado una transcripción automática del clip, se le ha puesto un título y también una serie de tags relacionados con la materia tratada.

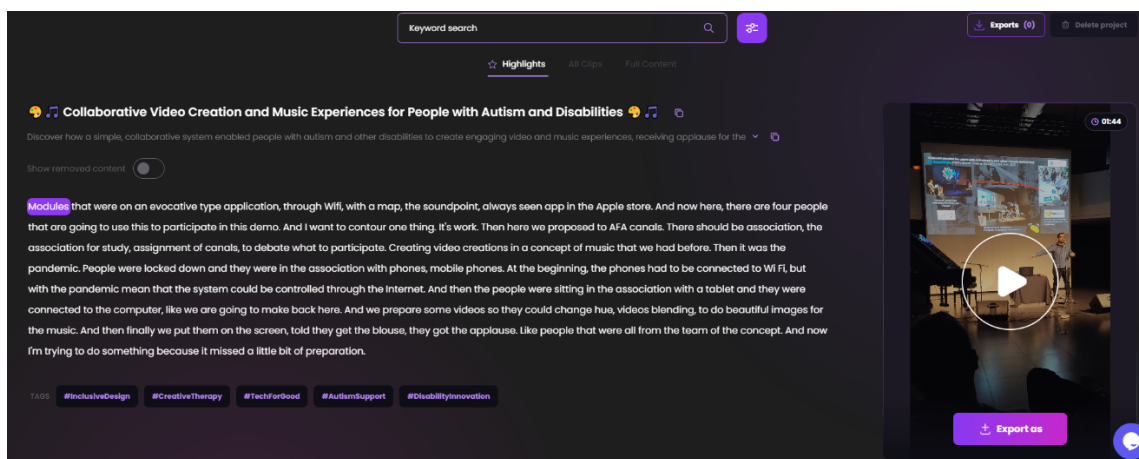


Ilustración 13. Vídeo procesado en Wisecut

Además, Wisecut ofrece una serie de clips diferentes entre los que se puede seleccionar el deseado, como se puede ver en la ilustración 12, cada uno con su título.

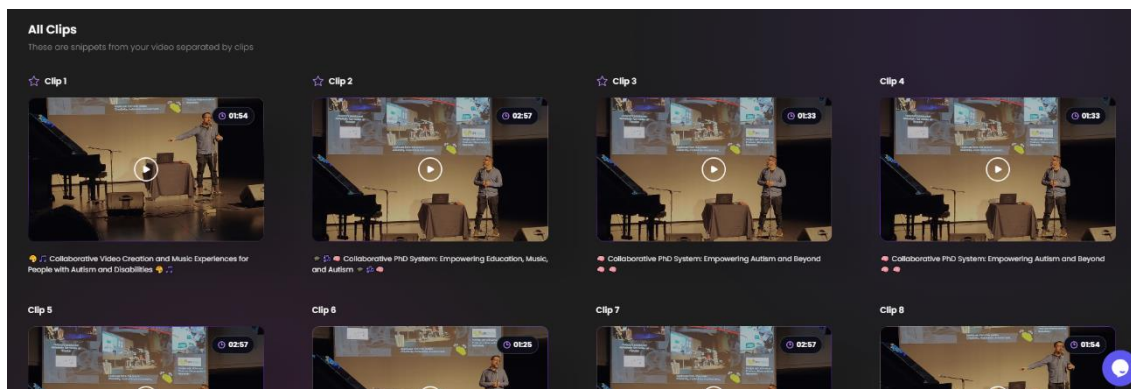


Ilustración 14. Selección de clips

Finalmente, también se genera una transcripción y rotulado del vídeo completo, sin editar, por si se desea utilizar.

Se ha optado por exportar el primer clip recomendado, y añadirle subtítulos.

Para finalizar, se utiliza Adobe Express para adaptar el vídeo al formato de Instagram Reels (vídeos cortos con información concisa) y su publicación en esta plataforma. Express ajusta el contenido al formato de vídeo utilizado por Instagram y dar los últimos toques de edición. Como se aprecia en la ilustración 13, se dispone de diferentes opciones para ajustar el tamaño al formato deseado.

Debido a que la resolución del vídeo era la misma que utiliza Reels, también se ha generado como alternativa una publicación normal de Instagram en su formato cuadrado, para que se aprecie la diferencia.

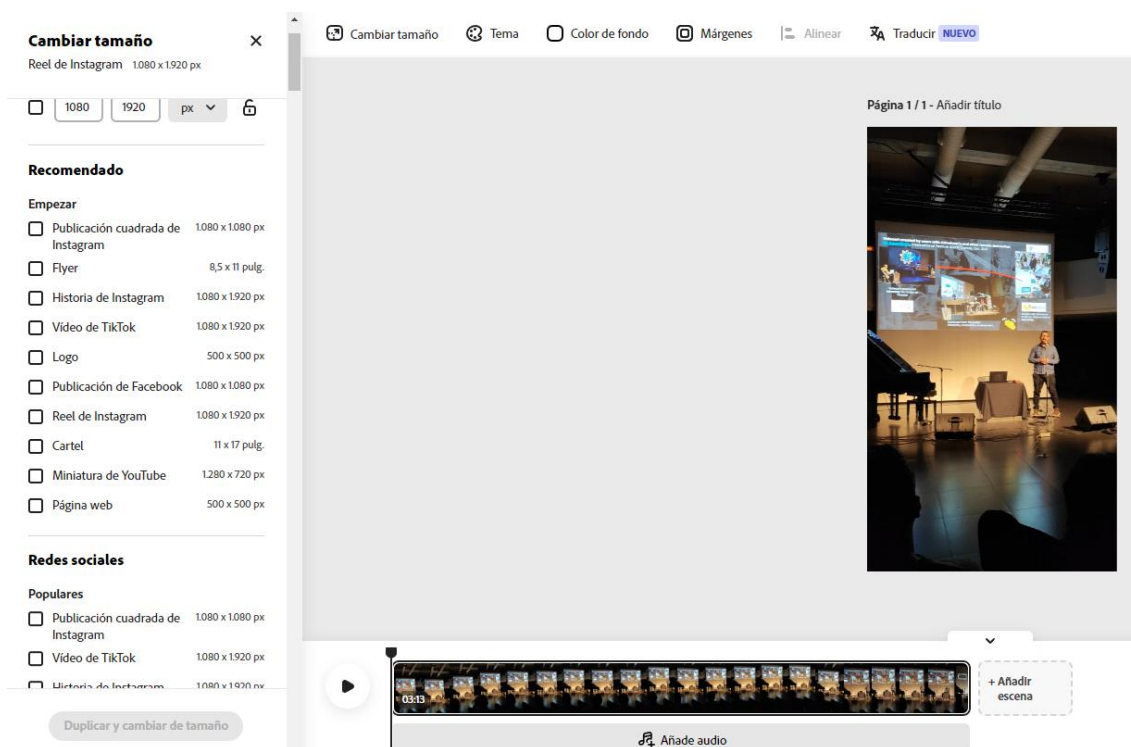


Ilustración 15. Express para Instagram Reels

Se puede apreciar también como utilizando esta misma aplicación se puede añadir una pista de audio deseada, de manera que, si se tiene el caso mencionado en el capítulo anterior en el cual se

desea integrar una pista de audio de mejor calidad para sustituir una grabada con una cámara u otro dispositivo, se puede utilizar Adobe Express para resolverlo

Al final, tras realizar estos pasos lo que se ha obtenido es un vídeo más corto, con una mejor calidad y volumen en la voz, sin ruido y, en este caso, adaptado al formato de vídeo que utiliza Instagram para Reels y para sus publicaciones normales.

Este caso concreto aborda Instagram, pero estas herramientas pueden utilizarse para publicar contenido en prácticamente cualquier red social respetando los límites y formatos de cada una.

Se añade a continuación un breve resumen del proceso:

1. Se ha editado el vídeo original con Enhance Speech para mejorar la nitidez de la voz y reducir ruido. El nivel de intensidad del efecto se ha fijado en 50% pero depende de cada caso, aún así, se recomienda nunca usar el 100% porque se genera un resultado antinatural. Se puede utilizar Clarity Vx como alternativa.
2. Se ha procesado el vídeo con menos ruido y mejor claridad de voz usando Wisecut para generar clips más cortos con los momentos más relevantes y para tener una transcripción y subtítulo del mismo. Como alternativa se puede utilizar Descript o Magisto.
3. Se ha elegido un clip de entre todos los sugeridos por Wisecut, se han corregido fragmentos que estaban mal subtítulos, se ha ajustado la duración del fragmento y se ha exportado y descargado.
4. El clip exportado de Wisecut se ha procesado con Adobe Express para ajustar su formato a la red social deseada, en este caso Instagram. Se sube el vídeo a la plataforma, se selecciona la opción de cambio de tamaño y en este caso se ha optado por un Instagram Reel y un post normal de Instagram. Como alternativa se podría utilizar Canva Pro.
5. Se ha obtenido, como resultado, un clip más corto que el original, con menos ruido, una voz más nítida, subtítulo y que se ajusta al formato de Instagram Reels y posts normales de Instagram.

En este ejemplo concreto no se ha utilizado la herramienta Emastered debido a que no había una pista de audio musical que requiriese una masterización como tal y la prioridad era la voz, siendo el resultado proporcionado por Enhance Speech de la calidad suficiente sin requerir tratamiento adicional. En un ejemplo en el cual se utilice Emastered se recomienda dejar las opciones y los parámetros en su configuración por defecto, debido a que modificarlas sin los conocimientos necesarios puede ocasionar resultados no deseados, aun así, se anima a cualquier persona en contacto con estas aplicaciones y herramientas a probar y testear cada una a modo de aprendizaje y así adaptarlas al oído, vista y gusto subjetivo de cada uno.

A continuación, se adjuntan enlaces de Google Drive con los diferentes resultados de los pasos seguidos en esta guía.

- Vídeo original sin editar: https://drive.google.com/file/d/1w6IMkGfw4SbuUFkeWp7J4D6G6340HujS/view?usp=drive_link
- Vídeo procesado con Enhance Speech para mejorar la calidad de la voz: https://drive.google.com/file/d/1DRScrpKCNxz_TIo8tP3IaKU3JErNhPL/view?usp=drive_link
- Vídeo procesado con Wisecut y subtítulo: https://drive.google.com/file/d/1-J1xqoe4ugo9EJt5lSnq5W8n8gKQICy/view?usp=drive_link
- Vídeo para Instagram Reels: https://drive.google.com/file/d/16E4dYyh4nnq0yHSd7siWA2kiRXfJdnA7/view?usp=drive_link
- Vídeo para post de Instagram: https://drive.google.com/file/d/1QUCzGQtEzLCIHAItNn04nBmLGIPSKLCq/view?usp=drive_link

Capítulo 9. Conclusiones y propuestas de trabajos futuros

En este trabajo se ha realizado una investigación en profundidad sobre la Inteligencia Artificial y sus aplicaciones enfocadas al tratamiento de contenido audiovisual, a lo largo de este proceso se han adquirido una serie de conocimientos sobre el mundo de la IA y sobre el alcance que, hoy en día, esta herramienta tiene en la vida de las personas.

Se plantean una serie de conclusiones obtenidas como parte del desarrollo de este proyecto:

- La Inteligencia Artificial está presente de una manera muy amplia y a veces poco obvia en el día a día de los seres humanos.
- Durante la preparación del proyecto se subestimó la cantidad de herramientas y aplicaciones que existen para los usos propuestos en este trabajo, existiendo un número inmenso de alternativas y opciones de software en diferentes formas que se usa para editar contenido audiovisual utilizando IA, así como para cualquier otro uso técnico.
- Se han cumplido y resueltos los objetivos planteados ya que se ha encontrado un proceso fácil, rápido y eficaz para editar contenido audiovisual utilizando Inteligencia Artificial y prepararlo para redes sociales.
- A pesar de la resolución propuesta, el proceso creativo audiovisual es muy subjetivo y personal, de manera que, aunque esta guía sirva como base, la enorme cantidad de recursos (incluso gratuitos) existentes en Internet y las posibilidades que estos ofrecen, dejan claro que hay muchísimas formas diferentes de haber resuelto los objetivos iniciales, por ejemplo, utilizando otras herramientas o priorizando otros aspectos del procedimiento.
- Se han revisado las políticas de privacidad de las herramientas utilizadas y la mayoría de ellas especifican de manera clara que los datos personales de usuario no se utilizan para entrenar sus sistemas (como por ejemplo Enhance Speech y Auphonic). Descript permite al usuario elegir cómo se utilizan sus datos y, como nota negativa, Wisecut y Emastered no especificaban este aspecto en sus políticas.

Se proponen, además, una serie de trabajos futuros que pueden utilizar este proyecto como base para desarrollar:

- Lo más básico sería analizar en mayor profundidad las herramientas que han sido descartadas por no ajustarse a los criterios buscados y utilizarlas como alternativa a la resolución planteada. Por ejemplo, Clarity Vx sería un plugin interesante que incorporar al proceso creativo aprovechando su potencia.
- Otra posibilidad es crear un proceso destinado a la edición audiovisual pero que sea un híbrido entre Inteligencia Artificial y herramientas más tradicionales, mezclando la edición tradicional manual y solo utilizando la IA para cosas puntuales, dejando así más libertad al usuario.
- Se plantea también enfocar las herramientas presentes en este proyecto a un uso más musical, por ejemplo, editar, mezclar y masterizar álbumes musicales de una manera más ágil que la habitual.
- Como propuesta que incluya Soundcool, se puede realizar un trabajo práctico en el cual editar el contenido en crudo del que disponga la plataforma utilizando estas herramientas y dejarlo listo para su publicación paulatina en redes sociales.
- Por último, dado que aparecen herramientas de IA cada día, sería interesante actualizar el estudio cada cierto tiempo para obtener las herramientas más ventajosas desde los diferentes puntos de vista expuestos.

9.1 Reflexión y valoración personal

Como final y cierre de este trabajo, he considerado interesante añadir una pequeña reflexión a título personal después de haber estado este tiempo en contacto con las tecnologías de Inteligencia Artificial, aprendiendo sobre ellas, comparándolas y finalmente usándolas para resolver los objetivos que tenía este proyecto.

Lo primero que me parece reseñable mencionar es que no era consciente de la cantidad de herramientas, páginas, software y demás a disposición del usuario, que se pueden utilizar casi sin ningún conocimiento previo y de manera gratuita, sobre todo a través de Internet. Era conocedor de las cosas habituales como ChatGPT y generadores de imágenes como Dall-E que fueron virales hace un par de años, pero eso es un porcentaje ínfimo de la totalidad de herramientas a las que cualquier persona puede acceder simplemente usando Google, ya sea para trabajo, ocio o creatividad. Hablando con personas para informarme sobre el mundo de las IA me ha quedado muy claro que muchísima gente, especialmente en nuestro campo, se apoya en estas herramientas para su trabajo diario y se han convertido en una parte más del día a día de mucha gente, esto sumado a la velocidad a la que avanzan cada día (por ejemplo la calidad de las imágenes generadas usando IA actual en comparación a las primeras de hace un par de años) hace que sea complicado prever el futuro y hasta qué punto son capaces de llegar, lo que está claro es que dentro de nuestro campo e incluso en otros no tan relacionados con la tecnología como podría parecer, se ha creado una necesidad de estar al día con los avances realizados porque por el contrario existe el riesgo de quedar desactualizado.

A título personal, como futuro graduado en esta Ingeniería y como Técnico de Sonido, probar estas herramientas me ha hecho darme cuenta de que son de una utilidad enorme para agilizar procesos que anteriormente resultaban laboriosos y costosos (masterizar audio, editar vídeo, ajustar volúmenes...), desde luego que el futuro es prometedor y la actualidad presenta muchas oportunidades de acelerar el trabajo y mejorarlo usando estas herramientas, pero hay una perspectiva un poco negativa relacionada con la capacidad de aprender de las IAs y que quizás en un futuro haga menos relevantes ciertas profesiones o ciertos trabajos realizados por humanos. Como todo en el campo de la innovación, pienso que lo mejor es mantener un equilibrio y utilizar estos conocimientos de una manera adecuada y cuidadosa.

Finalmente, cabe recalcar que, a pesar de que muchas herramientas de IA son bastante agresivas con los datos de usuario, las que he podido probar en este proyecto no utilizan en ningún caso los datos proporcionados para su entrenamiento y son, aparentemente, bastante respetuosas con lo que compartimos con ellas, lo cual ha sido una grata sorpresa.



Capítulo 10. Bibliografía

- [1] <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/> (consultado el 12 de agosto de 2024)
- [2] McCorduck P. “*Machines Who Think.*” 2004. Routledge editorial ISBN 9781568812052.
- [3] <https://www.forbes.com/sites/bernardmarr/2023/03/20/beyond-the-hype-what-you-really-need-to-know-about-ai-in-2023/> (consultado el 12 de agosto de 2024)
- [4] <https://pocketconfidant.com/how-the-era-of-artificial-intelligence-will-transform-society/> (consultado el 9 de agosto de 2024)
- [5] <https://podcast.adobe.com/enhance> (consultado el 1 de septiembre de 2024)
- [6] <https://www.adobe.com/es/express/> (consultado el 1 de septiembre de 2024)
- [7] <https://www.wisecut.ai/?lang=es> (consultado el 1 de septiembre de 2024)
- [8] García Talavera, Lucía. “*Soundcool: Smartphones, tablets y otras interfaces para la creación audiovisual.*” 2016.
- [9] <https://soundcool.org/> (consultado el 2 de septiembre de 2024)
- [10] <https://soundcool.org/proyectos/> (consultado el 2 de septiembre de 2024)
- [11] <https://www.youtube.com/watch?v=kkntxz4rRdw&t=826s> (consultado el 31 de agosto de 2024)
- [12] <https://developer.x.com/en/docs/x-api/v1/media/upload-media/uploading-media/media-best-practices> (consultado el 4 de septiembre de 2024)